arXiv:0902.2222v2 [hep-ex] 15 Aug 2009

The Extended-Track Event Reconstruction for MiniBooNE

R. B. Patterson [1 b], E. M. Laird [b], Y. Liu [a], P. D. Meyers [b], I. Stancu [a], H. A. Tanaka [2 b]

[a] *University of Alabama, Department of Physics and Astronomy, Tuscaloosa, AL, 35487 USA*

[b] *Princeton University, Department of Physics, Joseph Henry Laboratories, Princeton, NJ 08544 USA*

Abstract

The Booster Neutrino Experiment (MiniBooNE) searches for $\nu_\mu \rightarrow \nu_e$ oscillations using the $\sim$1 GeV neutrino beam produced by the FNAL Booster synchrotron. The array of photomultiplier tubes (PMTs) lining the MiniBooNE detector records Cherenkov and scintillation photons from the charged particles produced in neutrino interactions. We describe a maximum likelihood fitting algorithm used to reconstruct the basic properties (position, direction, energy) of these particles from the charges and times measured by the PMTs. The likelihoods returned from fitting an event to different particle hypotheses are used to categorize it as a signal ν_e event or as one of the background ν_μ processes, in particular charged current quasi-elastic scattering and neutral current π^0 production. The reconstruction and event selection techniques described here can be applied to current and future neutrino experiments using similar Cherenkov-based detection.

Key words:
PACS:

[1] Now at the California Institute of Technology, Physics Department 103-33, Pasadena, CA 91125, USA.

[2] Now at the University of British Columbia, Department of Physics and Astronomy, Vancouver, BC V6T 1Z1, Canada.

1 Introduction

Cherenkov detectors of $\gtrsim$1 kiloton have played a key role in the establishment of the phenomenon of neutrino oscillations [1a–b,2a–c,3a–c]. These detectors typically consist of a large volume of a homogeneous transparent medium (water or mineral oil) with a high index of refraction surrounded by an array of photomultiplier tubes (PMTs). Cherenkov photons produced by charged particles emerging from the neutrino interactions and traversing the medium with $\beta > 1/n$ (where n is the index of refraction and $\beta = v/c$) are detected by the PMTs. The photons are emitted at an angle θ_C relative to the track direction, where $\cos\theta_C = 1/n\beta$. The radiation is azimuthally symmetric about the track direction, resulting in a ring-like pattern that can be identified on the PMT array. The quantity, spatial distribution, and arrival times of these photons provide information on the location, direction, and energy of the particle. We refer to the extraction of such information from the charge and time measurements of the PMTs as "reconstruction." An analysis of the ring profile can also provide information on the identity of the particle and, if multiple rings are detected, the number of particles.

In this paper, we discuss the event reconstruction algorithms used in the Booster Neutrino Experiment (MiniBooNE), which searches for an excess of ν_e interactions indicative of $\nu_\mu \to \nu_e$ oscillations using the Fermilab $\sim$1 GeV neutrino beam. The MiniBooNE detector is a sphere of radius 610.6 cm filled with Marcol 7 mineral oil ($n \approx 1.47$) and divided into two optically isolated regions by an opaque shell of radius 575 cm, concentric with the sphere. The surface of the inner "main" region is instrumented with an array of 1280 inward-facing 8 in. PMTs which detect the light produced by the neutrino interactions. The outer "veto" region is instrumented with 240 PMTs and is used to tag charged particles that enter the detector from outside (*e.g.*, cosmic muons) or that exit the main region. Though no scintillator was added to the Marcol 7, it scintillates weakly, resulting in the production of delayed isotropic light for particles with sub- and super-Cherenkov velocities. A more detailed description of the experiment can be found in Refs. [4] and [5].

2 Neutrino Interactions at MiniBooNE

The $\sim$1 GeV neutrino beam at MiniBooNE results in interactions with relatively low outgoing particle multiplicity. The largest interaction channel is the charged current quasi-elastic (CCQE) process:

$$\nu_\ell + n \to \ell^- + p \tag{1}$$

which accounts for ∼40% of all neutrino interactions in the MiniBooNE detector. Since the recoil proton of Eq. (1) is typically below Cherenkov threshold, only the outgoing lepton produces significant light. In modeling such events within a reconstruction algorithm, one can therefore consider only the Cherenkov and scintillation light produced by the outgoing lepton. While the recoiling nucleon can produce scintillation light, this additional source of light is not considered in the reconstruction.

A muon in the MiniBooNE detector exhibits minimum-ionizing behavior through most of its path with little chance of radiative energy loss. In contrast, electrons typically induce electromagnetic showers, resulting in additional electrons and positrons that emit Cherenkov light. Thus, muons and electrons have significantly different Cherenkov ring patterns. This difference is the basis for distinguishing a muon track from an electron track, and the reconstruction algorithm has a model for each.

The next most abundant process in MiniBooNE is single pion production, which occurs primarily via Δ resonance or coherent scattering:

$$\begin{aligned} \nu_\ell + N &\rightarrow \ell^- + N' + \pi \qquad (\text{CC } \pi) \\ \nu_\ell + N &\rightarrow \nu_\ell + N' + \pi \qquad (\text{NC } \pi) \end{aligned} \tag{2}$$

where $N^{(\prime)}$ denotes a nucleon in the case of resonance production and a nucleus in the case of coherent production. Of particular interest is neutral current (NC) π^0 production. The two photons from $\pi^0 \rightarrow \gamma\gamma$ decay induce electromagnetic showers indistinguishable in MiniBooNE from those induced by electrons. Since the recoil nucleon/nucleus is typically below Cherenkov threshold, these NC π^0 events are well-described in the reconstruction algorithm by two electron-like tracks pointing back to a common vertex. The presence of two distinct Cherenkov rings allows for the separation of π^0 events from electron (ν_e CCQE) events. However, π^0 misidentification (for example, due to a large energy asymmetry between the two photons or due to a small photon opening angle which leaves the Cherenkov rings overlapping) accounts for most of the reducible background in the $\nu_\mu \rightarrow \nu_e$ oscillation search. Thus, successful reconstruction and identification of π^0 events is critical.

While other channels can be accommodated by the reconstruction algorithm, the ν_e appearance oscillation analysis uses only four event models: single electron track, single muon track, two γ tracks, and two γ tracks with a π^0 invariant mass.

3 The Detector Response

The Marcol 7 mineral oil used in MiniBooNE exhibits a rich array of optical phenomena despite its ∼20 m extinction length near the peak of PMT sensitivity (∼400 nm). Cherenkov and scintillation light production is accompanied by photon absorption, fluorescence (with several possible excitation/emission spectra and lifetimes), and Rayleigh and Raman scattering. The left plot in Figure 1 summarizes the rates of various processes as a function of wavelength. The cumulative extinction rate is shown as the black line. In the near-ultraviolet region (<320 nm), fluorescence processes dominate the extinction rate. The fluorescence lifetimes range from $1-35$ ns. In the visible region (>320 nm), the dominant processes are Rayleigh scattering and absorption.

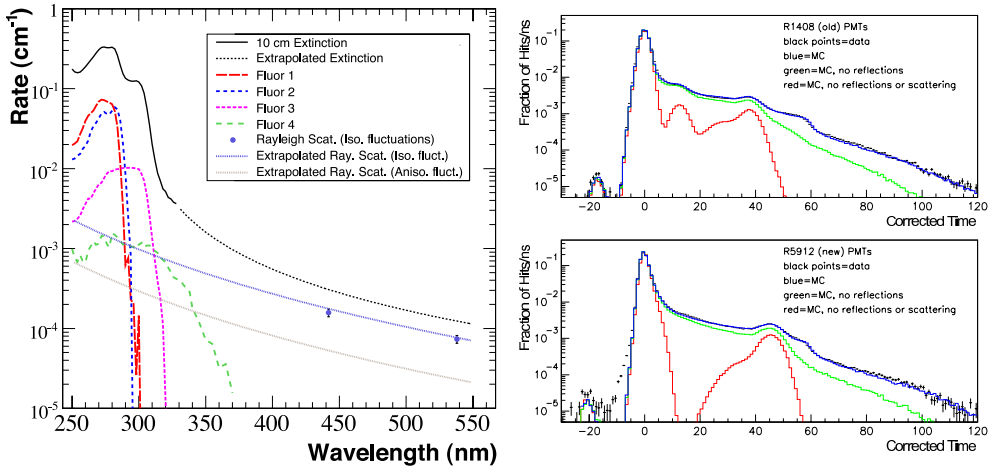

Fig. 1. Left: Rates of optical processes in Marcol 7 as a function of wavelength. The solid black line is the overall extinction rate obtained from laboratory measurements. The dashed black line is the extrapolated rate based on *in situ* data. The curves labeled "Fluor n" are the excitation rates for the four identified fluorescence processes. The light blue points represent the measured rates of Rayleigh scattering, and the dashed light blue and gray lines represent extrapolated rates. Right: Reconstructed photon arrival times for R1408 (top) and R5912 (bottom) PMTs for 397 nm light flashed from the center of the detector. The black histogram is the distribution from data and the blue is the complete Monte Carlo simulation. The green (red) shows the simulation with reflections (and scattering) turned off.

The 1520 8 in. PMTs in MiniBooNE are of two types: 322 model R5912 PMTs and 1198 model R1408 PMTs, both from Hamamatsu. All of the R5912 PMTs are located in the main PMT array. The PMT time response (particularly the late-pulsing behavior) and the variation of PMT efficiency with incident angle have been characterized in external measurements using a pulsed LED [6]. The PMTs *in situ* have been studied using a laser calibration system that produces sub-nanosecond bursts of nearly isotropic light in the detector.

The right panel of Figure 1 shows the times of PMT hits recorded in laser calibration events relative to their expected times. The colored histograms come from Monte Carlo simulation, with the red curve showing PMT and electronics timing effects (including pre- and late-pulsing), the green curve including photon scattering, and the blue curve adding photon reflections from the PMTs and the surface of the main detector region. The resulting time structure matches well with that seen in data (black points). We note that only the earliest photoelectron's time is reported when multiple photoelectrons are present in a hit. (A fresh hit can be seen after an electronics dead time of ∼300 ns.) The reported charge, however, does account for additional photoelectrons.

Absorption, scattering, reflections, and fluorescence processes influence both the topology and the time structure of hits. Additionally, scintillation light, emitted with a lifetime of ∼35 ns, adds intrinsically delayed light to the prompt Cherenkov photons. The event models within the reconstruction algorithm account for all these phenomena. In the discussion, we refer to Cherenkov and scintillation light that travels undisturbed from the particle track to the PMT as "direct" light. Light that undergoes fluorescence, scattering, or reflection we term "indirect" light.

A detailed GEANT3-based [7] Monte Carlo simulation of the MiniBooNE detector serves as the central tool for developing the reconstruction algorithm's predictive models. The accuracy of the models influences performance through resolution and biases in extracted track parameters and through the particle separation capabilities of the likelihood-based hypothesis testing. To balance the need for likelihoods that are as accurate as possible with the desire to contain the computational demands, detector and track model information extracted from the Monte Carlo simulation is tabulated, approximated, or parametrized wherever possible while minimizing any sacrifice in performance. Also, thanks to the excellent stability of the detector's optical properties, a single configuration of the reconstruction is sufficient for use on all MiniBooNE physics data.

4 The Single Track Model

The reconstruction of a single track in the detector, whether an electron or muon, is based on a model with seven parameters:

- the starting point: x_0, y_0, z_0
- the starting time: t_0
- the direction: θ_0, ϕ_0
- the kinetic energy: E_0.

The starting point and time are referred to as the "vertex" of the event. The direction is described in terms of the polar angle θ_0 defined with respect to the direction of the neutrino beam and the azimuthal angle ϕ_0 about this beam axis. For the electron model, the stochastic variations in the electromagnetic showers are modeled in average – no attempt is made to account for event-by-event variations in the shower profile. This is also true of the much smaller variations in the muon propagation (*e.g.*, straggling). This simplification allows the track to be fully specified from its initial properties summarized in the seven parameters above. We refer to this vector of parameters as $\mathbf{x}$.

The observables of an event are the measurements from the 1280 PMTs in the main region of the detector. For each PMT, we have:

- a bit indicating whether the tube registered a hit
- if the tube was hit, the measured charge of the hit
- if the tube was hit, the measured time of the hit.

Assuming that the PMTs behave independently, the likelihood for an observed set of PMT measurements given track parameters $\mathbf{x}$ can be expressed as:

$$\mathcal{L}(\mathbf{x}) = \prod_{\text{unhit}} (1 - P(i \text{ hit}; \mathbf{x})) \times \prod_{\text{hit}} P(i \text{ hit}; \mathbf{x})\ f_q(q_i; \mathbf{x})\ f_t(t_i; \mathbf{x}) \tag{3}$$

where the products are taken over all unhit and hit PMTs and:

- $P(i \text{ hit}; \mathbf{x})$ is the probability that the ith tube is hit, given parameters $\mathbf{x}$.
- q_i, t_i are the measured charge and time on the ith PMT.
- $f_q(q_i; \mathbf{x})$ is the probability distribution function (PDF) for the measured charge on the ith PMT, given $\mathbf{x}$, evaluated at q_i.
- $f_t(t_i; \mathbf{x})$ is the PDF for the measured time given $\mathbf{x}$, evaluated at t_i.

It is convenient to work with the negative logarithm of $\mathcal{L}$, and since the charge- and time-related portions of the likelihood are distinct, it is helpful to define:

$$F(\mathbf{x}) \equiv -\log \mathcal{L}(\mathbf{x}) \equiv F_q(\mathbf{x}) + F_t(\mathbf{x}) \tag{4}$$

where

$$\begin{aligned} F_q(\mathbf{x}) &= -\sum_{\text{unhit}} \log(1 - P(i \text{ hit}; \mathbf{x})) \ -\sum_{\text{hit}} \log\left(P(i \text{ hit}; \mathbf{x}) f_q(q_i; \mathbf{x})\right) \\ F_t(\mathbf{x}) &= -\sum_{\text{hit}} \log\left(f_t(t_i; \mathbf{x})\right) \end{aligned} \tag{5}$$

For brevity, we refer to F_q and F_t as the charge and time likelihoods, respectively, although they are the negative logarithms of the likelihoods.

If the number of photoelectrons n produced in a PMT is known, then $f_q(q)$ for the observed charge q is fully specified in terms of the response properties of the PMT (assumed fixed) without any reference to any other property of the detector. Further, n is a Poisson variable whose mean μ is a function of the track parameters $\mathbf{x}$. These considerations motivate a two-step approach to calculating the charge likelihood.

In the first step, using the known optical photon and particle propagation properties of the detector, one determines for a given particle type (electron/muon) and set of track parameters $\mathbf{x}$ the average number of photoelectrons that a particular PMT should observe. This quantity is referred to as μ_i, the "predicted charge" for the ith PMT. In calculating the predicted charge, one considers the quantity and angular distribution of light produced by the track, the absorption, scattering, and fluorescence of light as it traverses the mineral oil, the acceptance of the PMT, and anything else that influences the mapping $\mathbf{x} \mapsto \{\mu_i\}$.

For the second step, F_q given in Eq. (5) can be rewritten in terms of the predicted mean charges μ_i:

$$F_q(\mathbf{x}) \equiv -\sum_{\text{unhit}} \log(1 - P(i \text{ hit}; \mu_i)) - \sum_{\text{hit}} \log(P(i \text{ hit}; \mu_i) f_q(q_i; \mu_i)) \qquad (6)$$

where P and f_q are now functions of the $\{\mu_i\}$, with the mapping $\mathbf{x} \mapsto \{\mu_i\}$ from the first step left implicit. These two functions depend only on the properties of the PMT response and are, in particular, independent of the happenings within the detector volume. This stepwise approach decouples the track and optical photon model from the PMT response model. The latter is established once using laser calibration data. The laser system provides a source of approximately isotropic and prompt light with controllable intensity. The value of μ for a particular laser intensity can be determined from the occupancy of the PMTs. The function $f_q(q; \mu)$ for each PMT type can be determined by obtaining the distribution of measured charges q seen at various intensities μ. These distributions incorporate all processes associated with the charge response, such as the cascade of charge down the dynode chain, the amplification and digitization of the anode signal, and the conversion of the digital readouts into calibrated charge values. Examples of $f_q(q; \mu)$ obtained in this way are shown in Figure 2.

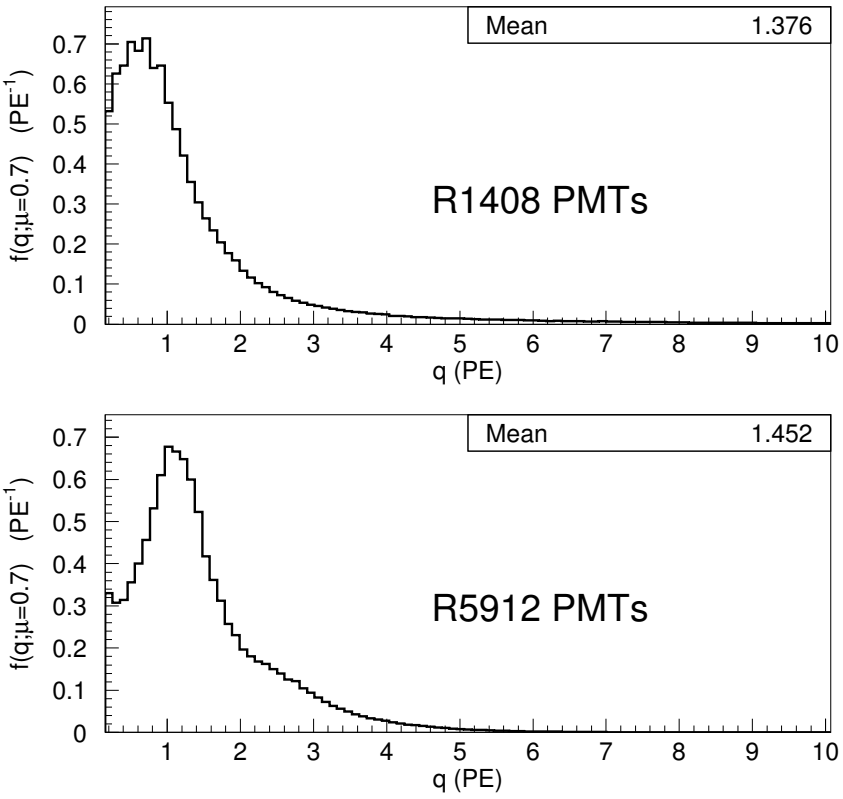

Fig. 2. Examples of $f(q;\mu)$ for $\mu = 0.7$ photoelectrons. Note that the distribution only includes cases where the PMT registers a hit.

4.2 Calculating the Predicted Charge μ

We begin our description of the predicted charge calculation with a simple scenario involving a fictitious point source of isotropic direct light. This is then generalized to a line source of isotropic light, appropriate for the scintillation production of a track. We then consider a line source in which the emission has a non-trivial (but azimuthally symmetric) angular distribution, representative of Cherenkov radiation. Finally, we account for the contribution of indirect light (*i.e.*, scattering, reflections, etc.) from both the scintillation and the Cherenkov emission. In order to simplify notation and without loss of generality, the discussion involves fixed but arbitrary PMT location and track parameters $\mathbf{x}$.

One needs only two parameters to describe the geometric relationship between an isotropic point source and a PMT. We choose (1) the distance r between the source and the PMT and (2) the angle of incidence η at the PMT, with $\eta = 0$ corresponding to incidence parallel to the PMT axis. The

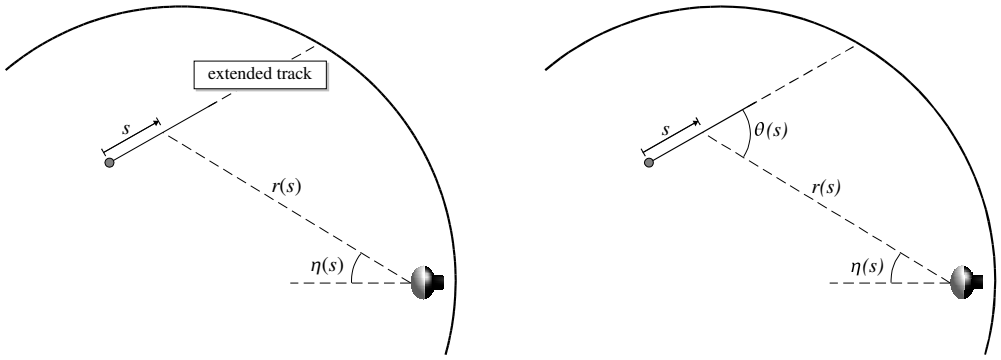

Fig. 3. Geometry of a line source of isotropic scintillation light (left) and directional Cherenkov light (right). In each case s is the distance from the origin of the track, $r(s)$ is the distance to the PMT from the point along the track, and $\eta(s)$ is the angle of PMT incidence for light emitted from this point. The presence of Cherenkov light requires the additional parameter $\theta(s)$ describing the light emission angle.

predicted charge on the PMT can be written

$$\mu_{\text{point,sci}} = \Phi_{\text{sci}}\ \Omega(r)\ T_{\text{sci}}(r)\ \epsilon(\eta) \tag{7}$$

where Φ_{sci} is an event-energy-dependent scintillation light yield (*i.e.*, $\Phi_{\text{sci}} = \Phi_{\text{sci}}(E_0)$), $\Omega(r)$ is a r-dependent solid angle factor, $T_{\text{sci}}(r)$ is the transmission of the oil and PMT glass as a function of light propagation distance, and $\epsilon(\eta)$ is the acceptance of the PMT as a function of the angle of incidence. The transmission depends on the wavelength spectrum of the light source and, thus, is specific in this case to scintillation. The wavelength-dependent PMT photocathode efficiency is also included in $T(r)$. The solid angle factor and PMT angular acceptance are purely geometric, independent of the particulars of the light production.

To generalize to an extended source of scintillation light, as depicted in Figure 3(left), the latter three factors in Eq. (7) become functions of s, the distance along the track from its origin. We must also include an emission profile $\rho_{\text{sci}}(s)$ describing the distribution of scintillation production along the track. This profile satisfies $\int_{-\infty}^{\infty} \rho(s) ds = 1$. The predicted charge can now be obtained by integrating along the track's axis:

$$\mu_{\text{sci}} = \Phi_{\text{sci}} \int_{-\infty}^{\infty} ds\ \rho_{\text{sci}}(s)\, \Omega(s)\, T_{\text{sci}}(s)\, \epsilon(s)\ . \tag{8}$$

The dependences of Ω, T_{sci}, and ϵ on r and η have been recast as dependences on s, as the relationship $s \mapsto (r, \eta)$ is fixed for given track parameters and PMT location. Figures 4 and 5 show $\rho_{\text{sci}}(s)$ for 300 MeV muons and electrons, obtained via Monte Carlo simulation. The distribution is relatively flat for muons until the end of the track, where the ionization rate per unit track length

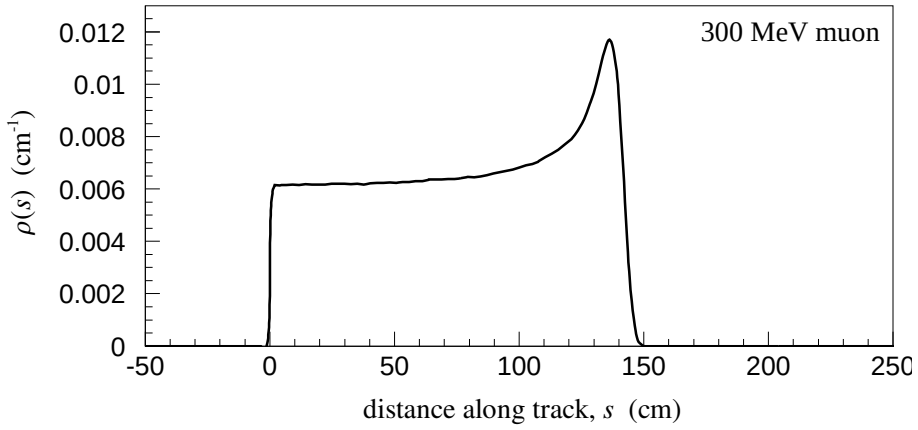

Fig. 4. Scintillation emission profile for 300 MeV muons as a function of s, the distance along the track.

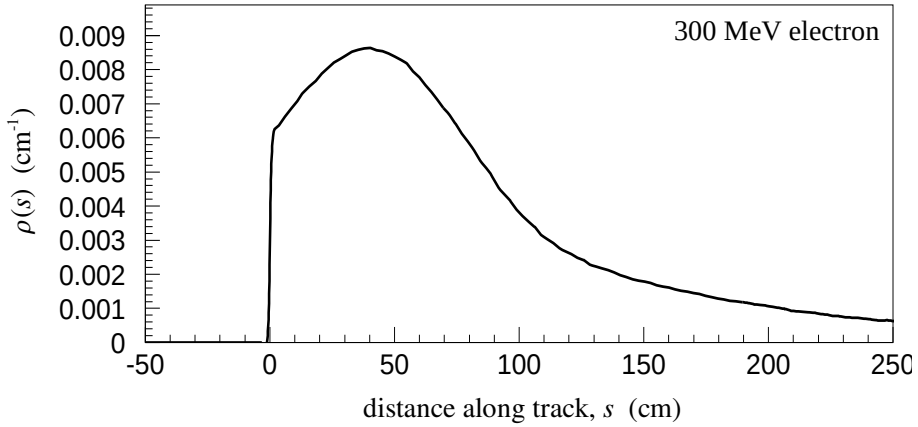

Fig. 5. Scintillation emission profile for 300 MeV electrons as a function of s, the distance along the track.

becomes larger. (Saturation effects are taken into account in the simulation.) For electrons, the distribution reflects the showering behavior of the track, with $\rho_{\rm sci}(s)$ rising and falling with the number of shower particles.

We now consider an extended track emitting light with a non-trivial angular distribution, as is the case for Cherenkov radiation. We introduce θ, the angle to the PMT from the track, as shown in Figure 3(right), and express the predicted charge $\mu_{\rm Ch}$ due to Cherenkov radiation as:

$$\mu_{\rm Ch} = \Phi_{\rm Ch} \int_{-\infty}^{\infty} ds \rho_{\rm Ch}(s)\, \Omega(s)\, T_{\rm Ch}(s)\, \epsilon(s)\, g(\cos\theta(s); s) \tag{9}$$

This expression differs from its scintillation counterpart by the presence of an angular emission profile $g(\cos\theta(s); s)$. Note that g depends on s in two distinct ways: the angular profile of the light changes as the track propagates and loses energy, and the angle θ to the PMT changes depending on which

part of the track we are considering.

Figures 6 and 7 show $\rho_{\rm Ch}(s)$ and $g(\cos\theta; s)$ for simulated 300 MeV muons and electrons. As the muon propagates, loses energy, and approaches the Cherenkov threshold, the rate of Cherenkov radiation per unit track length decreases and the Cherenkov angle becomes smaller. The scattering of the muon, which causes deviations of the track from its original direction, has been included in the Monte Carlo simulation and results in the spread of the angular distribution about the nominal $\cos\theta_C$. For electrons, $\rho_{\rm Ch}(s)$ follows the shower profile, like $\rho_{\rm sci}(s)$. The presence of the shower particles is readily apparent in the $g(\cos\theta; s)$ distribution, which becomes substantially wider as the "track" propagates.

In writing Eq. (9), a few simplifications are made. A more rigorous treatment would involve $\frac{d^2P(\cos\theta,\phi)}{d\cos\theta d\phi}$, the differential probability of sending an emitted Cherenkov photon in the direction $(\cos\theta, \phi)$, which would be integrated over the solid angle subtended by the PMT. By assuming that $\frac{d^2P(\cos\theta,\phi)}{d\cos\theta d\phi}$ is constant over the PMT face, the integration can be effected by simply multiplying the differential probability by the solid angle Ω. Also, the azimuthal symmetry of the emission reduces the two-dimensional PDF to a one-dimensional expression, namely $g(\cos\theta)$. We take $g(\cos\theta)$ to satisfy

$$\int_{-1}^{1} d\cos\theta\, g(\cos\theta; s) = 1 \tag{10}$$

for all values of s, with the result that a factor of $(4\pi)^{-1}$ is absorbed into the definition $\Phi_{\rm Ch}$.

4.3 Indirect Light (scattering, fluorescence, reflections)

The above formalism determines the predicted charge for light arriving at the PMTs directly from the track without any redirection. However, the detector has sources of indirect light from scattering, fluorescence, and reflection as discussed in Section 3.

The geometry for indirect light given scintillation (isotropic) emission is shown in Figure 8(left), where an infinitesimal element ds along the track is situated at radius R from the center of the detector and at angle Θ relative to the position of the PMT. The direct light from this track element is simply the integrand of Eq. (8):

$$d\mu_{\rm sci}^{\rm direct} = ds\ \Phi_{\rm sci}\, \rho_{\rm sci}(s)\, \Omega(s)\, T_{\rm sci}(s)\, \epsilon(s)\ . \tag{11}$$

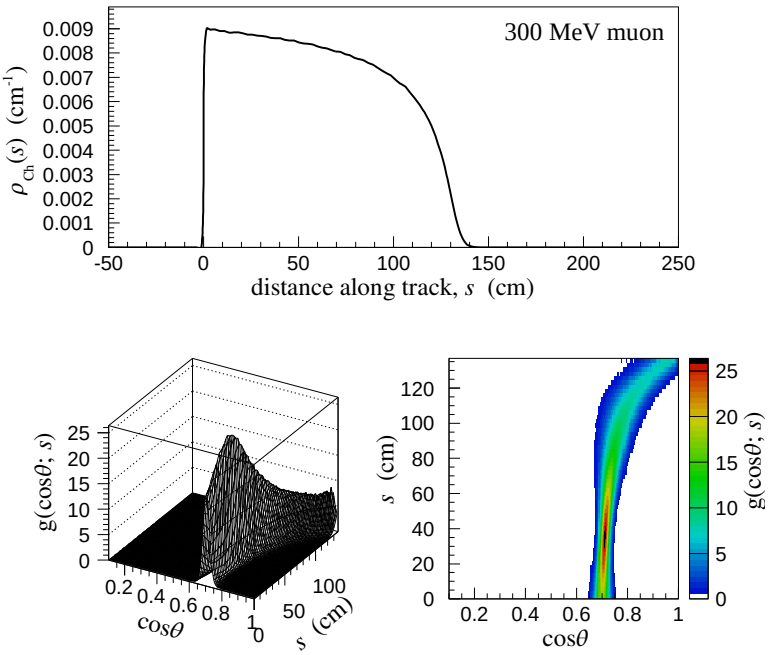

Fig. 6. Top: Cherenkov emission profile for 300 MeV muons as a function of s, the distance along the track. Bottom: Two-dimensional emission profile showing the angular distribution of Cherenkov light as a function of s for a 300 MeV muon.

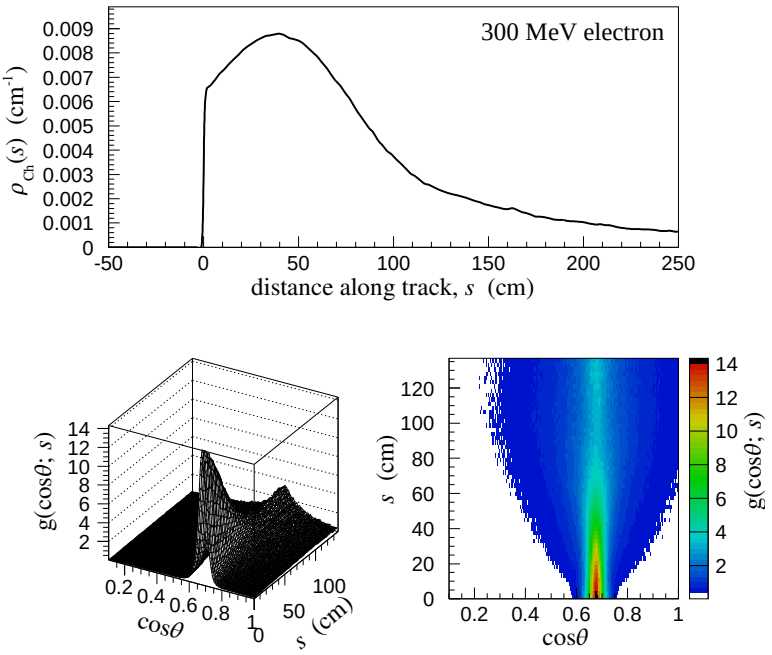

Fig. 7. Top: Cherenkov emission profile for 300 MeV electrons as a function of s, the distance along the track. Bottom: Two-dimensional emission profile showing the angular distribution of Cherenkov light as a function of s for a 300 MeV electron.

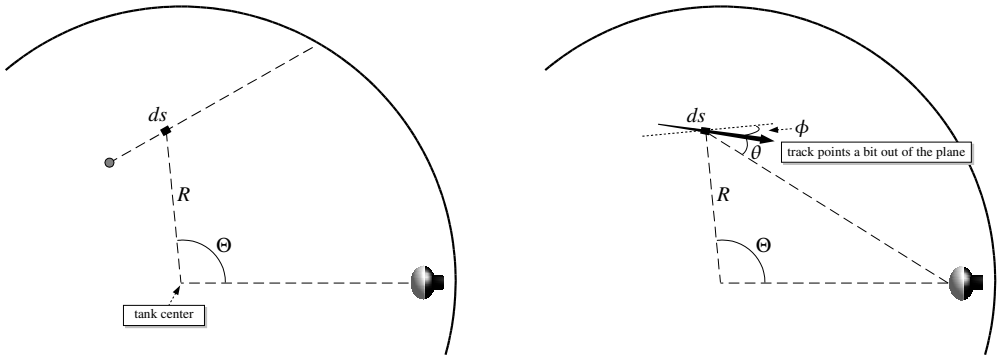

Fig. 8. Geometry of indirect light from an isotropic (left) and directional (right) source of light.

An analytic expression for the *indirect light* would involve an elaborate integral over emission angles and scattering points throughout the tank. Rather than attempt this, we observe that the value of such an integral must be proportional to the source strength and must otherwise depend only on the topological variables R and Θ. The source strength dependence can be eliminated by forming a ratio of the indirect and direct light predictions:

$$A_{\rm sci}(R, \cos\Theta) \equiv \frac{d\mu_{\rm sci}^{\rm indirect}}{d\mu_{\rm sci}^{\rm direct}} \ . \tag{12}$$

$A_{\rm sci}$, which we refer to as the *scattering table* and which we build via the detector simulation, is a property only of the detector optics and the $(R, \cos\Theta)$ of the track element. With this table, the indirect light contribution can be immediately incorporated into the expression for predicted charge:

$$\mu_{\rm sci} = \Phi_{\rm sci} \int_{-\infty}^{\infty} ds\, \rho_{\rm sci}(s)\, \Omega(s) T_{\rm sci}(s) \epsilon(s) \left[1 + A_{\rm sci}\left(R(s), \cos\Theta(s)\right)\right] \tag{13}$$

where the dependence of R and Θ on s has been made explicit.

For Cherenkov light, the situation is more complex since the light emission is anisotropic. Two additional variables are needed to specify the direction of a *vector* source relative to the PMT position and the tank center. We make the following non-unique choice for these two variables, as shown in Figure 8(right):

- θ: the angle between (1) the source direction vector and (2) the source-to-PMT ray (the same θ as defined elsewhere).
- ϕ: The angle between (1) the plane containing the tank center, the PMT, and the source, and (2) the plane containing the track and the tank center.

The intensity of the indirect Cherenkov light is normalized with respect

to a fictitious *isotropic* Cherenkov source with predicted charge $d\mu_{\rm Ch}^{\rm direct,iso}$, so that:

$$A_{\rm Ch}(R, \cos\Theta, \cos\theta, \phi) \equiv \frac{d\mu_{\rm Ch}^{\rm indirect}}{d\mu_{\rm Ch}^{\rm direct,iso}} \tag{14}$$

The total Cherenkov contribution to the mean predicted charge, including both direct and indirect components is

$$\mu_{\rm Ch} = \mu_{\rm Ch}^{\rm direct} + \mu_{\rm Ch}^{\rm indirect} \tag{15}$$

where $\mu_{\rm Ch}^{\rm direct}$ is given by Eq. (9) and $\mu_{\rm Ch}^{\rm indirect}$ is given by

$$\begin{aligned}\mu_{\rm Ch}^{\rm indirect} = \Phi_{\rm Ch} \int_{-\infty}^{\infty} ds \, [&\rho_{\rm Ch}(s)\, \Omega(s)\, T_{\rm Ch}(s)\, \epsilon(s) \\ &\times A_{\rm Ch}\left(R(s), \cos\Theta(s), \cos\theta(s), \phi(s)\right)]\end{aligned} \tag{16}$$

Implicit in this expression is the assumption that the Cherenkov angular emission profile does not change as the track propagates (and loses energy). This simplification has negligible impact, as the processes that result in indirect light tend to destroy the initial emission pattern anyway.

We now have a complete charge prediction, where the total predicted charge for the PMT is given by:

$$\mu = \mu_{\rm Ch} + \mu_{\rm sci} \tag{17}$$

4.4 Computation of Integrals

The predicted charge integrals above would be impractical to evaluate with sufficient spatial granularity within a maximum likelihood fit, where they must be evaluated multiple times for every PMT in every event. To mitigate this, all integrations are performed and tabulated beforehand as follows.

The integrand of Eq. (8), which pertains to direct scintillation light, is the product of a "source" factor $\Phi_{\rm sci}\rho_{\rm sci}(s)$ and an "acceptance" factor

$$J(s) \equiv \Omega(s) T_{\rm sci}(s) \epsilon(s) \ . \tag{18}$$

As in Eq. (8), the dependence on r and η is recast through s. If $J(s)$ varies

gradually enough, it can be approximated with a parabolic form:

$$J(s) = j_0 + j_1\, s + j_2\, s^2 \ . \tag{19}$$

An example of this is shown in Figure 9. The predicted charge then becomes

$$\mu_{\rm sci}^{\rm direct} = \Phi_{\rm sci} \left[j_0 \int_{-\infty}^{\infty} ds\ \rho_{\rm sci}(s) + j_1 \int_{-\infty}^{\infty} ds\ \rho_{\rm sci}(s)\, s + j_2 \int_{-\infty}^{\infty} ds\ \rho_{\rm sci}(s)\, s^2 \right] \ . \tag{20}$$

The first integral is identically unity. The remaining two depend on the energy E_0 via $\rho_{\rm sci}(s)$, but they depend on no other track parameters. This allows one to tabulate the integrals beforehand as a function of E_0, eliminating their evaluation from the actual likelihood calculation.

The parabolic coefficients $\{j_i\}$ are obtained by evaluating Eq. (18) at three points along the track: the start of the track ($s=0$), the midpoint of the Cherenkov emission profile ($s=\Delta s_{\rm mid}$) and twice this distance from the start ($s=2\Delta s_{\rm mid}$). The choice of where to evaluate $J(s)$ is somewhat arbitrary; one could choose any three points that sample a large fraction of the range over which light is produced. The extension to indirect isotropic light is straightforward with a redefinition of $J(s)$ that incorporates the scattering table:

$$J_{\rm sci}^{\rm indirect}(s) \equiv \Omega(s)\, T_{\rm sci}(s)\, \epsilon(s)\, A_{\rm sci}\left(R(s), \cos\Theta(s)\right) \ . \tag{21}$$

An analogous three-integral expression can be formed for the Cherenkov predicted charge:

$$\begin{aligned} \mu_{\rm Ch}^{\rm direct} = \Phi_{\rm Ch} \Bigg[& j_0 \int_{-\infty}^{\infty} ds\ \rho_{\rm Ch}(s) g(\cos\theta(s); s) \\ & + j_1 \int_{-\infty}^{\infty} ds\ \rho_{\rm Ch}(s)\, g(\cos\theta(s); s) s \\ & + j_2 \int_{-\infty}^{\infty} ds\ \rho_{\rm Ch}(s)\, g(\cos\theta(s); s) s^2 \Bigg] \ . \end{aligned} \tag{22}$$

As in the scintillation case, the integrals depend on the energy of the track. However, they also depend on two parameters which define the PMT-track geometry. We choose these parameters to be the vertex-to-PMT distance, $r(0)$; and the cosine of the angle-to-PMT viewed from the vertex, $\cos\theta(0)$. Labeling the integrals $\mathcal{I}_i^{\rm Ch}$, we obtain

$$\mu_{\rm Ch}^{\rm direct} = \Phi_{\rm Ch} \left(j_0 \mathcal{I}_0^{\rm Ch} + j_1 \mathcal{I}_1^{\rm Ch} + j_2 \mathcal{I}_2^{\rm Ch} \right) \ . \tag{23}$$

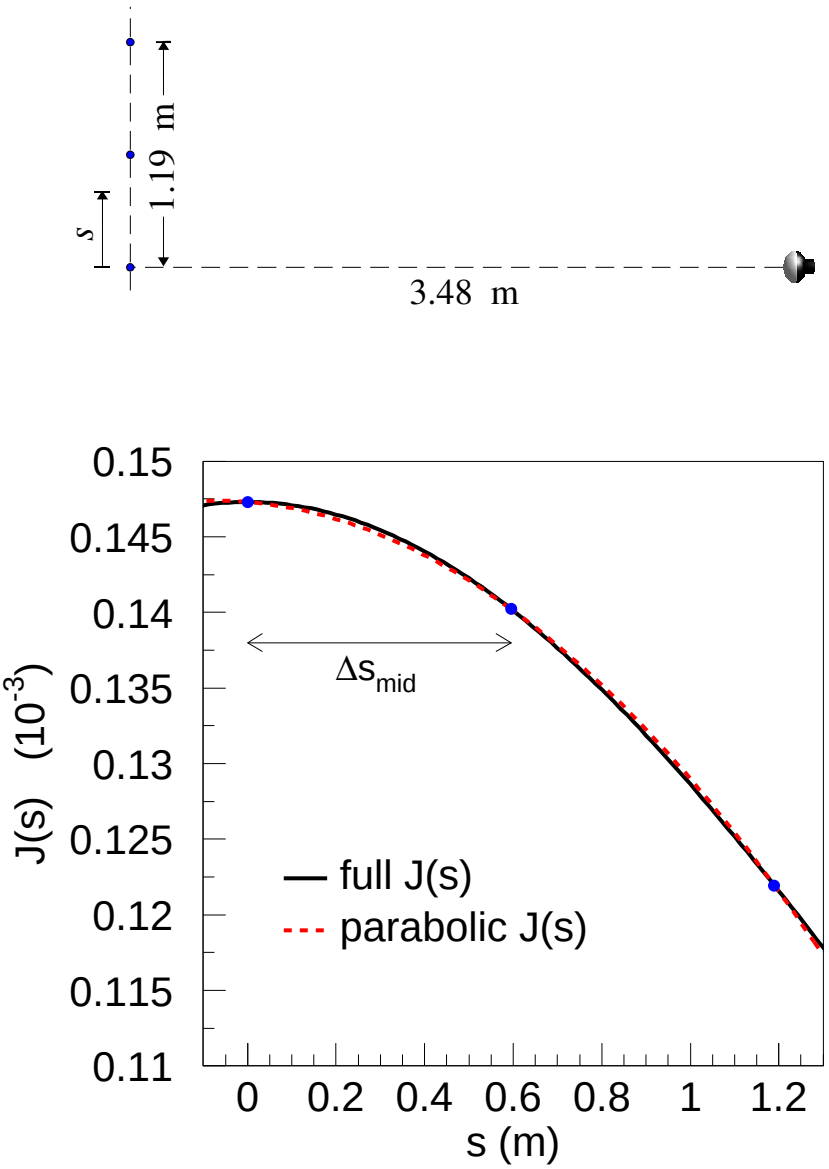

Fig. 9. Example of $J(s)$ for the track geometry shown in the top figure, where a muon 3.48 meters from a PMT moves perpendicular to the PMT axis. $J(s)$ is evaluated at three points located along the track, corresponding to the three points in the figure. In the bottom panel, the parabolic approximation for $J(s)$ based on the three evaluated points (dashed red) is compared to the exact form (solid black).

Although the integrals $\left\{I_i^{\rm Ch}\right\}$ depend on three parameters, it is still feasible to tabulate their values ahead of time, eliminating their explicit evaluation within the likelihood calculation.

For indirect Cherenkov light, recall that the scattering table $A_{\rm Ch}$ is normalized to a fictitious isotropic Cherenkov source. Thus, the parabolic method

used for indirect scintillation can be applied here to give

$$J_{\rm Ch}^{\rm indirect}(s) \equiv \Omega(s)\, T_{\rm Ch}(s)\, \epsilon(s)\ A_{\rm Ch}\left(R(s), \cos\Theta(s)\right)\ , \tag{24}$$

with two energy-dependent integrals $I_{i=1,2}^{\rm Ch} = \int ds \rho_{\rm Ch}(s) s^i$ to be tabulated.

4.5 The Time Likelihood

The time portion of the likelihood involves the PDF $f_t(t; \mathbf{x})$ for the measured PMT hit time t given track parameters $\mathbf{x}$. Some of the dependence on the track parameters can be eliminated by using the "corrected time" t^c:

$$t^c = t - t_0 - \frac{r(\Delta s_{\rm mid}(E_0))}{c_n} - \frac{\Delta s_{\rm mid}(E_0)}{c}\ . \tag{25}$$

The expression removes from t three terms, namely the starting time of the track t_0, the expected time for light to propagate from the track midpoint to the PMT (with the speed of light in mineral oil given by c_n), and the time for the particle to propagate from its starting point to the midpoint. The dependence of $\Delta s_{\rm mid}$ on the track energy E_0 is made explicit. Defining t^c with respect to the track midpoint (as opposed to, say, the start of the track) improves the validity of the simplifications we make below.

Given the $t \mapsto t^c$ substitution, we seek the PDF $f_{t^c}(t_i^c; \mathbf{x})$ for the corrected time given arbitrary PMT-track configurations. The space of configurations is five dimensional. However, producing tables of $f_{t^c}(t^c)$ as a function of five parameters is impractical. To reduce the task, we make the assumption that the corrected time PDF depends only on the track energy, the predicted "prompt" charge, and the predicted "late" charge, where "prompt" corresponds roughly to light arriving at the PMT directly from the track without delays induced by emission lifetimes, scattering, *etc.* Loosely, the assumption is that the shape of the corrected time spectrum is dominated by the physical extent of the track (characterized by its energy), and by the amount of prompt and late light reaching the PMT. The extent of the track affects the spread of possible hit times, since there is a spread of photon production times, while the amounts of prompt and late light affect the proportion of peaked "prompt" response to the tail of "late" response.

This assumption reduces the configuration to three dimensions. We make one further simplification by assuming that for a given energy E_0, $f_{t^c}(t^c)$ can be modeled by separate primitive prompt and late distributions, indexed by $\mu_{\rm prompt}$ and $\mu_{\rm late}$, respectively, from which we can construct the full PDF. As a result, the PDF is indexed not by the fundamentally two-dimensional space

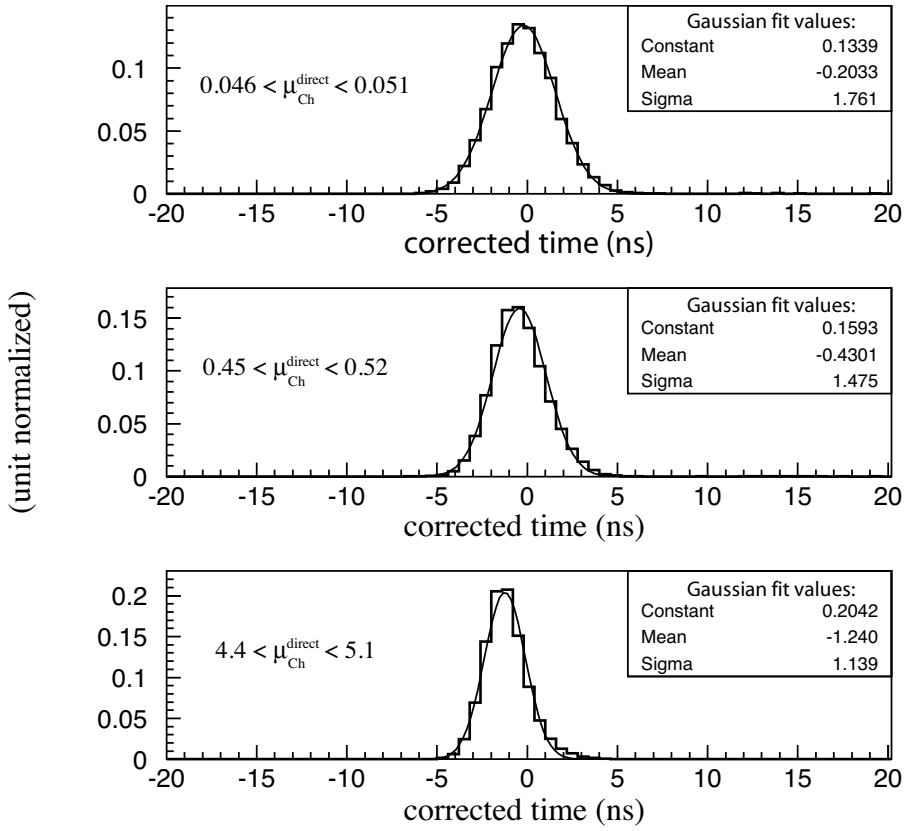

Fig. 10. Distributions of t^c for direct Cherenkov light for R5912 PMTs. The primary difference between the two PMT types is that the R1480 PMTs show wider prompt timing peaks than do the newer R5912 PMTs.

of $(\mu_{\text{prompt}}, \mu_{\text{late}})$, but by μ_{prompt} and μ_{late} separately, with the full PDF calculated on-the-fly as described below. In practice, Cherenkov and scintillation primitive distributions are created and used as proxies for the desired prompt and late distributions, respectively.

The Cherenkov (prompt) primitive distributions are created by simulating particles throughout the detector with isotropically chosen directions and fixed energy E_0. The particles are created with direct Cherenkov light only; all other sources of light (scintillation, scattering, etc.) are turned off. For each event, the true track parameters are used to evaluate the direct Cherenkov predicted charge $\mu_{\text{Ch}}^{\text{direct}}$ for each hit. Histograms of the corrected time t^c of the hits are produced for various ranges of predicted charge. Figure 10 shows three such t^c distributions for 300 MeV muons. Since only direct Cherenkov light is present, the histograms show no late-time features. The shapes of the time spectra depend on the $\mu_{\text{Ch}}^{\text{direct}}$ values involved, with the distributions becoming narrower and earlier with increasing predicted charge due to the increasing probability that an early photon will be recorded at the PMT. The t^c distribution in each predicted charge range is fit to a Gaussian distribution, and the resulting Gaussian parameters (mean and width) are subsequently parametrized across

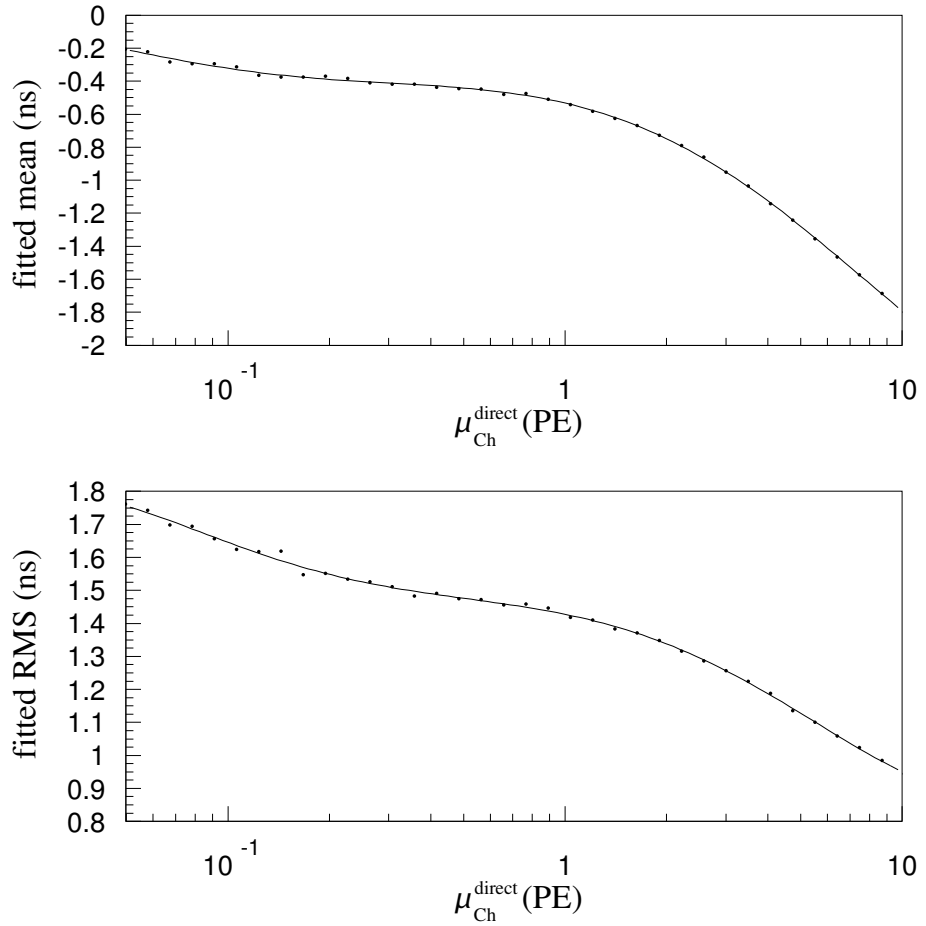

Fig. 11. For 300 MeV muons, the mean and width parameters from Gaussian fits like those in Figure 10 versus $\mu_{\rm Ch}^{\rm direct}$. The dependence is parameterized as a sixth-order polynomial in $\log(\mu_{\rm Ch}^{\rm direct})$.

$\mu_{\rm Ch}^{\rm direct}$ using a sixth-order polynomial in $\log(\mu_{\rm Ch}^{\rm direct})$. Figure 11 shows an example of these "second-level" fits for 300 MeV muons. The procedure is repeated at many values of E_0, with the two second-level fits providing seven parameters for the Gaussian mean and seven parameters for the Gaussian width at each energy. The energy dependences of these 14 parameters are then fit as a function of E_0 to fourth-order polynomials in a third-level parametrization, as shown in Figure 12 for the first seven of the second-level parameters. In addition to conveniently summarizing the dependence of the time response on $\mu_{\rm Ch}^{\rm direct}$ and E_0, the parametrizations provide a smooth likelihood surface, as required by the minimization algorithm.

This completes the prompt primitive distributions $G_{\rm Ch}(t^c; E_0, \mu_{\rm Ch}^{\rm direct})$. They are calculated for a given predicted charge $\mu_{\rm Ch}^{\rm direct}$ at a PMT for a track with energy E_0 by evaluating the 14 third-level curves which parametrize the E_0 dependence of the second-level functions. These 14 values then determine the two functions which parametrize the dependence of the mean and width of the

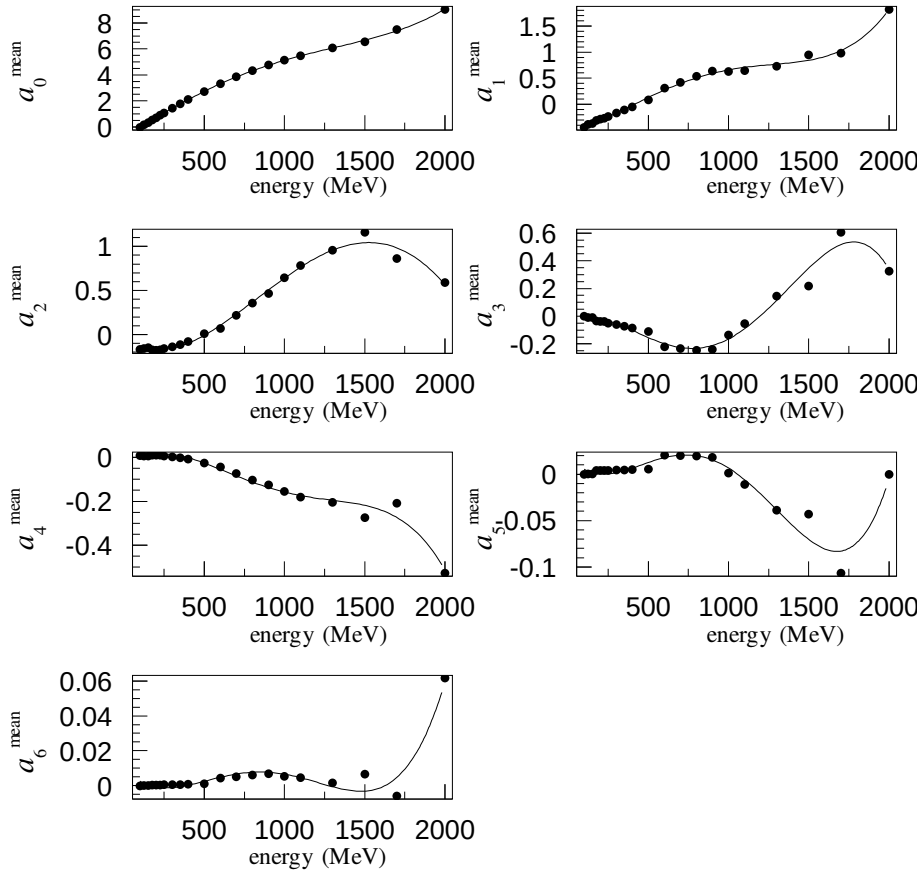

Fig. 12. Third-level fits parametrizing the energy dependence of the second-level parameters describing the Gaussian mean and width parameters.

Gaussian parameters on $\mu_{\text{Ch}}^{\text{direct}}$ at E_0. These are evaluated at the particular value of $\mu_{\text{Ch}}^{\text{direct}}$ to determine the appropriate mean and width of the Gaussian distribution which describes the Cherenkov primitive distribution appropriate for the particular values of $\mu_{\text{Ch}}^{\text{direct}}$ and E_0. The validity of the assumptions and simplifications can be tested directly by comparing the calculated primitive distributions with those actually observed in the Monte Carlo simulated events. Figure 13 shows such a comparison.

For the scintillation (late) primitive distributions, events are generated with scintillation light only, and the t^c histograms are created in bins of $\mu_{\text{sci}}^{\text{direct}}$. Each t^c histogram is fit to a sum of two exponentials convolved with a Gaussian distribution. The exponential decay constants are fixed at $\tau_1 = 5\,\text{ns}$ and $\tau_2 = 30\,\text{ns}$, leaving three free parameters: the time origin, the Gaussian resolution, and the relative weight of the two exponentials. At a given E_0, the three parameters are extracted in each histogram to obtain their dependence on $\mu_{\text{sci}}^{\text{direct}}$, which is parametrized as a sixth-order polynomial in $\log(\mu_{\text{sci}}^{\text{direct}})$. The dependence of the three sets of seven parameters from the polynomials are parametrized as a function of E_0 by fourth-order polynomials, in analogy with the Cherenkov case. We now have $G_{\text{sci}}(t^c; E_0, \mu_{\text{sci}}^{\text{direct}})$, the scintillation primitive

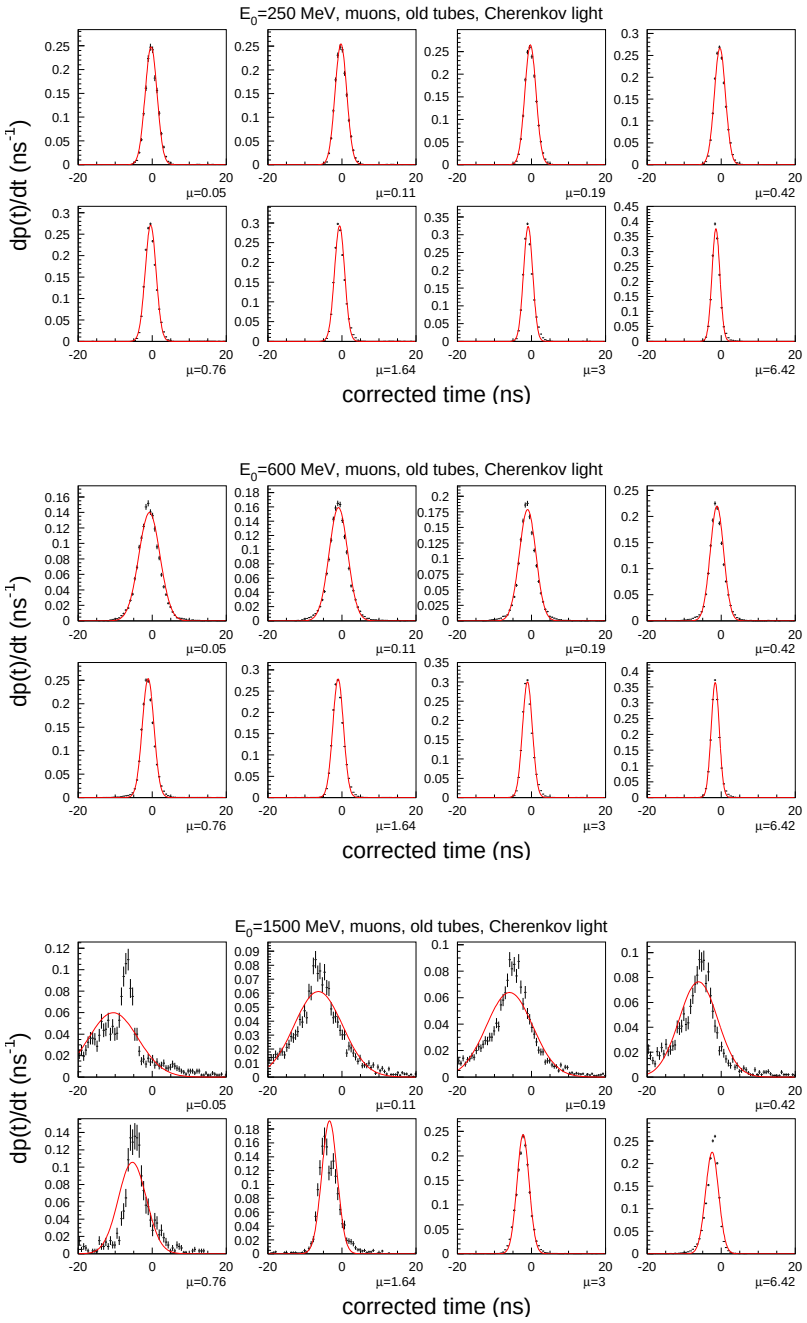

Fig. 13. Check of primitive distributions. The parametrized t^c likelihood distributions for Cherenkov light are compared against the actual t^c distributions for 250 (top), 600 (middle) and 1500 (bottom) MeV muons as a function of predicted charge. (Note that the high-E_0, low-μ hits, which are less well modeled, are rare.)

distribution. Figure 14 shows the primitive distributions compared with actual t^c distributions from the Monte Carlo simulation for muons at $E_0 = 250$, 600 and 1500 MeV.

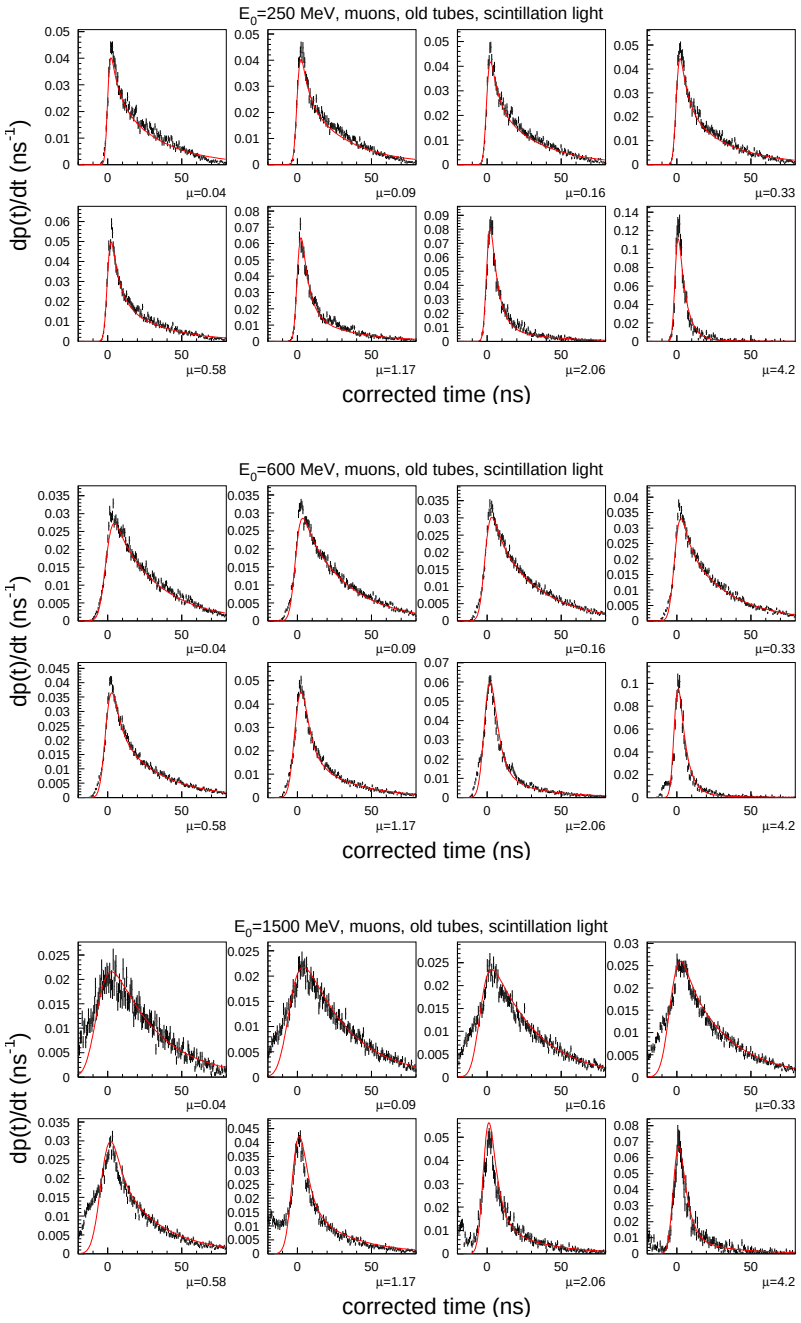

Fig. 14. Check of scintillation primitive distributions. The parametrized t^c likelihood distributions for scintillation light are compared against the actual t^c distributions for 250 (top), 600 (middle) and 1500 (bottom) MeV muons as a function of predicted charge.

To obtain the complete PDF $f_{t^c}(t^c)$ for a given PMT hit, we first divide the predicted charge into "prompt" and "late" components based on the sources:

$$\begin{aligned} \mu_{\text{prompt}} &= 0.95\mu_{\text{Ch}}^{\text{direct}} \\ \mu_{\text{late}} &= 0.05\mu_{\text{Ch}}^{\text{direct}} + \mu_{\text{Ch}}^{\text{indirect}} + \mu_{\text{sci}}^{\text{direct}} + \mu_{\text{sci}}^{\text{indirect}} \end{aligned} \tag{26}$$

These definitions of μ_{prompt} and μ_{late} encapsulate the following assumptions. First, all prompt light is due to direct Cherenkov light; all other light (including indirect Cherenkov light) follows the late scintillation-based time distribution. The 5% of direct Cherenkov light which is included in μ_{late} accounts for the late pulsing of PMTs which causes promptly arriving light to appear at late times. The late time distribution is used to model indirect Cherenkov light since the fluorescence, scattering, and reflection processes which lead to indirect light have a time structure similar to that of scintillation. Further, late timing is less critical to the reconstruction, as prompt light provides essentially all the useful timing information.

The μ_{prompt} and μ_{late} values are used to combine the prompt and late primitive distributions to form the total PDF $f_{t^c}(t^c)$. In the process, one must account for the fact that a PMT can register only one hit for a given track due to the latency period of the electronics. Since the time reported by the PMT reflects the time of the first photoelectron, we assume that a PMT hit that includes a prompt photoelectron obeys the prompt primitive distribution regardless of how many late photoelectrons may follow. The Poisson distribution gives us the probabilities that no prompt or late photoelectrons are produced based on their expected mean values:

$$\begin{aligned} P(\text{no prompt PEs}) &= e^{-\mu_{\text{prompt}}} \\ P(\text{no late PEs}) &= e^{-\mu_{\text{late}}} \end{aligned} \tag{27}$$

With these probabilities, one can calculate the probability that a given hit contains at least one prompt photoelectron:

$$P(\text{prompt PE present}|\text{hit}) = \frac{1 - P(\text{no prompt PEs})}{1 - P(\text{no prompt PEs})P(\text{no late PEs})} \tag{28}$$

We assign this as the weight w_p for the prompt primitive distribution within the total PDF, while $w_l = 1 - w_p$ is used as the weight for the late primitive distribution, yielding:

$$f_{t^c}(t^c; E_0, \mu_{\text{prompt}}, \mu_{\text{late}}) = w_p G_{\text{Ch}}(t^c; E_0, \mu_{\text{prompt}}) + w_l G_{\text{sci}}(t^c; E_0, \mu_{\text{late}}) \tag{29}$$

where G_{Ch} and G_{sci} are the prompt Cherenkov-based and late scintillation-based timing distributions, respectively. Note that Eq. (28) removes the overall probability of a hit occurring. Thus, even if the absolute probability of a prompt photoelectron is small, $w_p \sim 1$ if $\mu_{\text{late}} \ll \mu_{\text{prompt}}$.

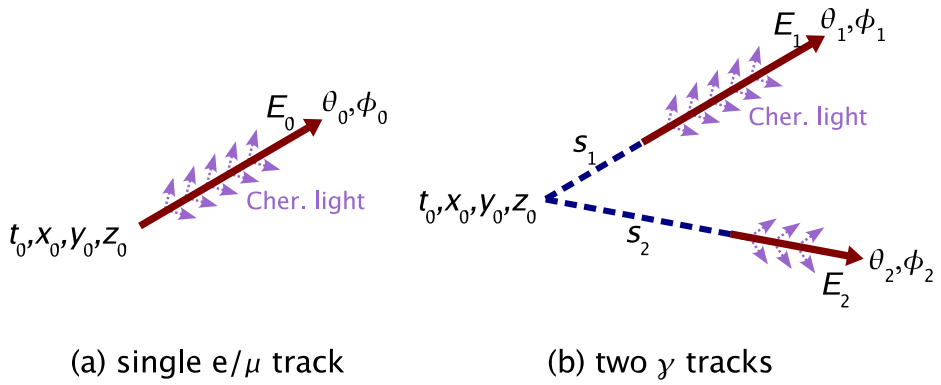

Fig. 15. Internal fit parameters for (a) a single muon or electron track and (b) two photon tracks. Each photon track includes a conversion distance parameter s. The two-track parameters can be constrained such that the invariant mass of the two photons $M_{\gamma\gamma}$ is always M_{π^0}.

5 The Two-track Reconstruction

The single track model with seven parameters is sufficient for reconstructing ν_μ and ν_e CCQE events as well as cosmic muons or their decay electrons. As discussed in Section 2, NC π^0 events require a two-track model. Figure 15 shows the 12 parameters needed to describe two photon tracks originating from a common vertex. The electron track model serves double-duty as the photon track model, under the assumption that a showering electron is indistinguishable from a showering photon apart from the conversion distance. (The mean conversion length in Marcol 7 is 67 cm.) While the two-track model is conceptually a straightforward extension of the one-track case, complexities of the likelihood space require special treatment in the likelihood maximization algorithm.

5.1 Two-track Charge Likelihood

Since the charge likelihood F_q depends only on the measured charge q and the total predicted charge μ at each PMT, the two-track charge likelihood can be obtained by adding together the predicted charges from the two tracks to obtain the total predicted charge at each PMT:

$$\mu = \mu_{\text{track 1}} + \mu_{\text{track 2}} \tag{30}$$

The one-track time likelihoods account for the "first photoelectron only" nature of the electronics by calculating the probability that a prompt photoelectron is present and by then weighting the prompt primitive distribution's influence on the full time PDF accordingly. This scheme is extended to handle two tracks as follows.

The two single-track primitive distributions, $G^i_{\rm Ch}$ and $G^i_{\rm sci}$ are formed for each track, where i labels the track number. In anticipation of aggregating all late light later, the late "scintillation" primitive distributions are averaged to form:

$$\hat{G}_{\rm sci} = \frac{1}{2}\left(G^1_{\rm sci} + G^2_{\rm sci}\right) \tag{31}$$

Of the two tracks, one will have a midpoint that is nearer to the target PMT than the other. This track's quantities are labeled with "n" (near); the other's, "f" (far). In analogy with Eq. (26), we define:

$$\begin{aligned}
\mu_{\rm prompt,n} &\equiv 0.95\ \mu^{\rm direct}_{\rm Ch,n} \\
\mu_{\rm prompt,f} &\equiv 0.95\ \mu^{\rm direct}_{\rm Ch,f} \\
\mu_{\rm late} &\equiv \mu_{\rm tot} - \mu_{\rm prompt,n} - \mu_{\rm prompt,f}
\end{aligned} \tag{32}$$

The probabilities of not obtaining a photoelectron from these sources are:

$$\begin{aligned}
\overline{P}_{\rm n} &\equiv P(\text{no prompt PE from near track}) &&= e^{-\mu_{\rm prompt,n}} \\
\overline{P}_{\rm f} &\equiv P(\text{no prompt PE from far track}) &&= e^{-\mu_{\rm prompt,f}} \\
\overline{P}_{\rm l} &\equiv P(\text{no late PEs}) &&= e^{-\mu_{\rm late}}
\end{aligned} \tag{33}$$

from which the following weights are obtained:

$$\begin{aligned}
w_{\rm n} &= \frac{1-\overline{P}_{\rm n}}{1-\overline{P}_{\rm n}\overline{P}_{\rm f}\overline{P}_{\rm l}} \\
w_{\rm f} &= \frac{1-\overline{P}_{\rm f}}{1-\overline{P}_{\rm f}\overline{P}_{\rm l}}(1-w_{\rm n}) \\
w_{\rm l} &= 1 - w_{\rm n} - w_{\rm l}
\end{aligned} \tag{34}$$

where $w_{\rm n}$ is the probability that a prompt photoelectron from the near track exists given that any photoelectron exists, *etc.* The weights are used to combine the two prompt primitive distributions and the averaged late primitive distribution:

$$f_{t^c}(t^c) = w_{\rm n}\, G_{\rm Ch,n}(t^c) + w_{\rm f}\, G_{\rm Ch,f}(t^c) + w_{\rm l}\, \hat{G}_{\rm sci}(t^c) \tag{35}$$

resulting in the complete two-track corrected time PDF.

6 Minimization of $-\log(\mathcal{L})$

6.1 One-track

With the likelihood $\mathcal{L}$ defined, the parameter set $\mathbf{x}$ that minimizes its negative logarithm $F \equiv -\log(\mathcal{L})$ must be found. In the single track minimization process, two issues arise.

First, the energy parameter E_0 is tied to the geometry of the event via the track profiles $\rho(s; E_0)$. If the spatial parameters (x_0, etc) are varied simultaneously with E_0, the minimization algorithm can get confused by this correlation. The solution is to iterate the minimization process, as described below.

A second issue arises from the discrete PMT lattice, which imprints a small fluctuating signal on $\mathcal{L}(\mathbf{x})$ despite the smooth parameterizations of the input tables. Minimization algorithms that rely on gradients or that do not have controllable step sizes are thus poorly suited to the problem. Therefore, the SIMPLEX method in Minuit[11] is chosen over the MIGRAD method.

To initiate the minimization process, a vector of seed parameters is obtained from a fast fitter. In the first iteration of the minimization, the energy E_0 is held at its seeded value while the six remaining parameters (x_0, y_0, z_0, t_0, θ_0, ϕ_0) are varied via Minuit/SIMPLEX to find a temporary minimum of F. These six parameters are then fixed, while E_0 is freed for a second iteration. Finally, E_0 is fixed once more with the six other parameters varied (as in the initial iteration), to find the final minimum of F. The parameters from the final SIMPLEX call are returned as the best-fit set $\mathbf{x}$.

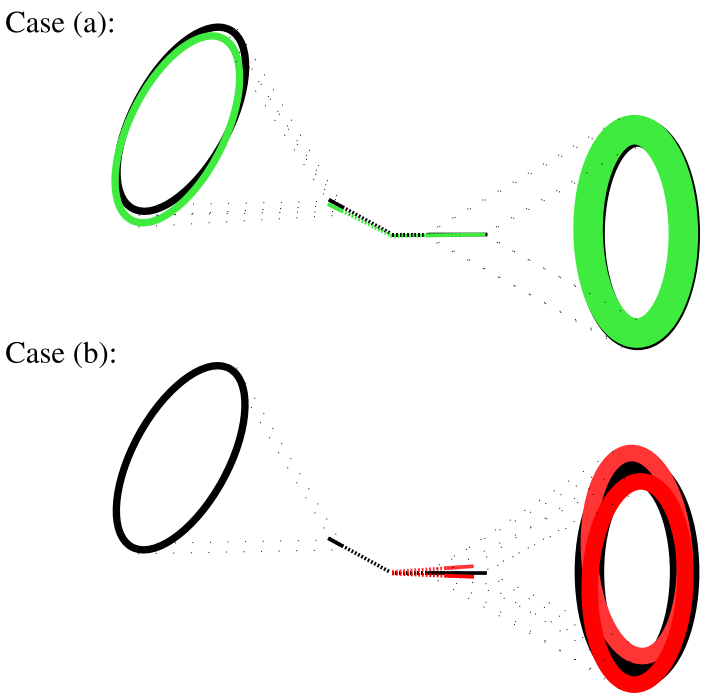

Fig. 16. The two π^0 fitter configurations discussed in the text.

6.2 Two-track

Figure 16 shows two situations that may arise in fitting a π^0 event. The event has two photon tracks whose Cherenkov rings are represented in black, while green and red represent two ring configurations corresponding to different starting values for the parameters. In case (a), the 12 parameters are near the correct configuration. In case (b), the parameters are such that both tracks are directed toward the dominant ring. The latter case corresponds to a local minimum which may offer a worse likelihood than case (a); however, the minimization algorithm may be trapped by the fact that any small changes to the parameters result in an increased value of F. For example, sweeping one track over to the smaller black ring involves passing through a region with little detected Cherenkov light. The intermediate states form a barrier of disfavored values of F that the minimization algorithm code will have difficulty overcoming, especially if several parameters must be adjusted simultaneously even to make the distant configuration an improvement.

The two-track fits require a minimization approach that avoids these traps in the likelihood surface. Monte Carlo π^0 events were used to identify both common and rare trapping scenarios, and the minimization algorithm addresses each of these through the following collection of seed configurations and minimization sequences:n

- The conversion lengths s_1 and s_2 are seeded with either 50 cm or 250 cm, leading to four possible pairs.
- θ_1 and ϕ_1 are seeded with the results from the electron single-track fit or with one of the eight perturbations (see below).
- θ_2 and ϕ_2 are seeded with the "best" directions from a full grid of tested directions (see below).
- The four-vertex of the event is seeded with the four-vertex from the electron single-track fit shifted according to s_1, θ_1 and ϕ_1.
- E_1 is seeded with approximately E_0 from the electron single-track fit, while the seed for E_2 is, when possible, the energy needed to give an invariant mass equal to M_{π^0}.

For the last item, the energy seeds are based on an empirical expression that accounts for the second ring's energy contribution to the one-track fit energy as a function of the angle between the two γ tracks. The function is obtained by simulating NC π^0 events and reconstructing them with the single track electron fit. Using the true photon energies ($E_1 > E_2$) and opening angle $\theta_{\gamma\gamma}$, the relationship between $(E_0 - E_1)/E_2$ and $\cos\theta_{\gamma\gamma}$ is parametrized by a linear function. This function is used to determine the seed E_1 and E_2 values for a given E_0 and $\cos\theta_{\gamma\gamma}$.

The nine (1+8) possible seed directions for track 1 are created as follows. The spatial distribution of PMT charge is projected onto a plane perpendicular to the electron fit direction. The major axis of the covariance ellipse of the resulting two-dimensional distribution is found. The nine possible seed directions are rotations of the best electron-fit direction by 0, $\pm$0.159, $\pm$0.450 and $\pm$0.644 radians parallel to the major axis and $\pm$0.159 radians perpendicular to the major axis.

The seed directions for track 2 come from either a "coarse" 24×12 (θ, ϕ) grid or a "fine" 50×25 (θ, ϕ) grid. The grid type is specified at run time.

Of the 9×(24×12)=2592 or 9×(50×25)=11250 possible combinations of track 1 and track 2 seed directions, only six are used to seed actual minimization sequences. They are the six combinations with (1) the best total likelihood (time and charge), (2) the best charge likelihood when the estimated track energies[3] are similar ($0.5 < \frac{E_1}{E_2} < 2$), (3) the best charge likelihood when the estimated track energies are dissimilar, and (4)−(6) the second-best likelihood in each of these three cases. A small exception: combinations involving one of the eight perturbed track 1 directions *and* having estimated track energies that differ by more than a factor of five are not considered. In these energetically asymmetric cases, the electron fit direction would be a good representation of track 1 since the influence of track 2 would be small, so there is no reason to

[3] Recall the energy seed procedure discussed above.

consider alternate track 1 directions.

The six best and second-best track 1 and track 2 direction and energy combinations are chosen anew for each permutation of (s_1, s_2), leaving 24 complete parameter sets in total. Each set seeds two Minuit sequences. In sequence 1, the SIMPLEX minimization routine is first run with the energy and angle parameters held fixed; once a minimum is found, SIMPLEX is called again with all parameters free. In sequence 2, only the energy parameters are held fixed in the initial SIMPLEX call. This results in 48 fit sequences, each of which ends with a completely free SIMPLEX minimization. The best parameter set (*i.e.*, the one that results in the lowest F) from the forty-eight final SIMPLEX calls is reported as the best parameter set $\mathbf{x}$ for the two-track fit. During all minimizations, the photon conversion points are constrained to be within the main detector volume. This prevents the fit from removing the influence of a track through an inflated conversion length. The event vertex itself, however, is not constrained.

The SIMPLEX minimizations can be performed with a constraint on the invariant mass. This is accomplished by removing E_2 as a free parameter and setting it via the relation:

$$E_2 = \frac{M_{\pi^0}^2}{2E_1(1-\cos\theta_{\gamma\gamma})} \tag{36}$$

where $\theta_{\gamma\gamma}$ is the angle between the photon tracks. This fixed-mass mode is the actual π^0 hypothesis, while the free-mass mode allows for mass reconstruction. Both are used for π^0 identification, with the former lending its maximized likelihood $\mathcal{L}_{\pi^0}$ and the latter providing a reconstructed mass $M_{\gamma\gamma}$.

7 Performance

The performance of the reconstruction algorithm on simulated neutrino interactions (ν_μ CCQE, ν_e CCQE and NC π^0 events) generated with NUANCE [8] and propagated through the detector Monte Carlo simulation is shown in Figures 17−21.[4] The electron and muon one-track fits are applied to the ν_μ and ν_e CCQE events, respectively, while the two-track fits, with and without the mass constraint, are applied to the NC π^0 events. A minimal selection is applied to ensure that the events are contained within the main

[4] For comparisons of Monte Carlo to data and for discussion of the development and validation of the Monte Carlo, see Refs. [9] and [10].

detector region and, for plots not directly related to vertex reconstruction, within the fiducial volume of the analysis.

A few observations about the figures, particularly Figure 18: (1) The resolutions for ν_μ CCQE events degrade as the energy approaches the Cherenkov threshold from above, as most of the information in the event comes from the (rapidly decreasing) prompt Cherenkov light. (2) A low-energy π^0 offers little information on its direction, as all that is observed is the isotropically produced $\gamma\gamma$ final state. Thus, the angular resolution (but not the vertex resolution) becomes poor. (3) The vertex resolution for π^0 events is worse than ν_e CCQE events since the former has spatially displaced light production thanks to the $\sim$70 cm average photon conversion length. (4) Typical performance numbers for ν_e CCQE events: 20 cm vertex resolution, 12% kinetic energy resolution, and 4° angular resolution. (5) Typical performance numbers for ν_μ CCQE events: 10 cm vertex resolution, 8% kinetic energy resolution and 2° angular resolution. (6) Figure 19 shows a π^0 mass resolution of $\sim$13%.

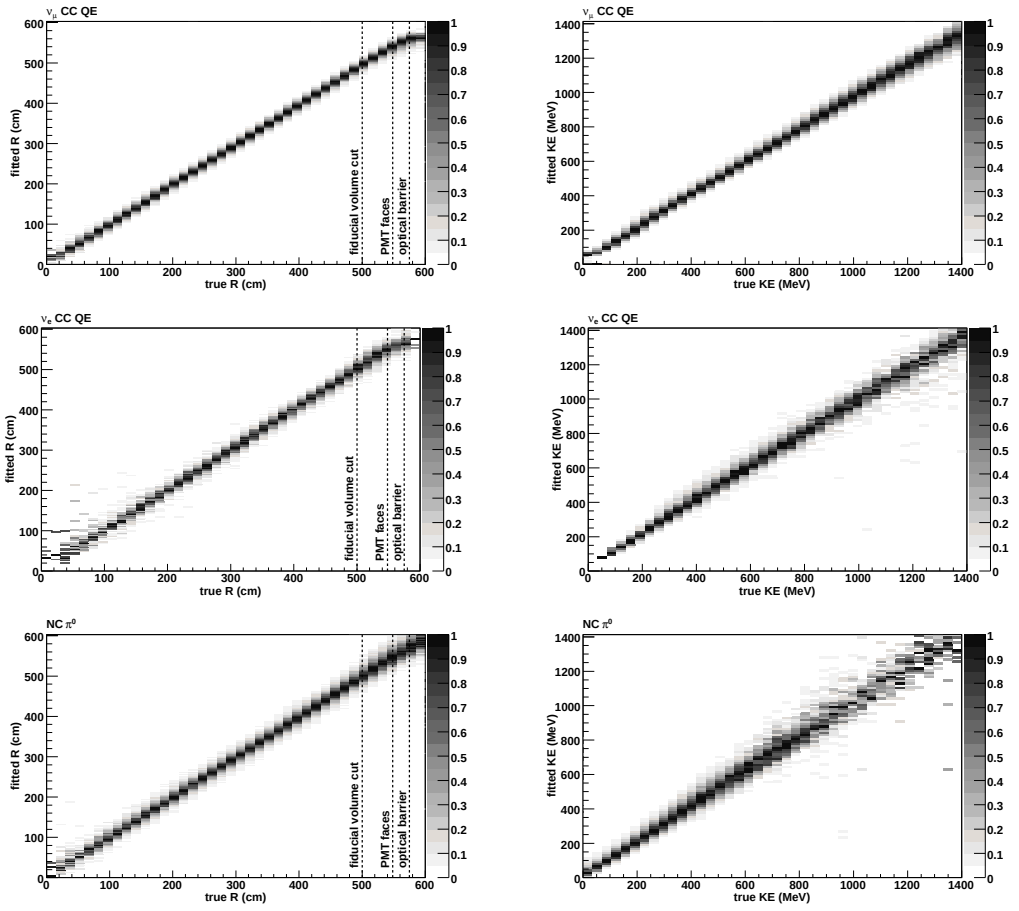

Fig. 17. Reconstructed radial event position (left) and kinetic energy (right) plotted against their true values. The top, middle, and bottom panels show ν_μ CCQE, ν_e CCQE, and NC π^0 events.

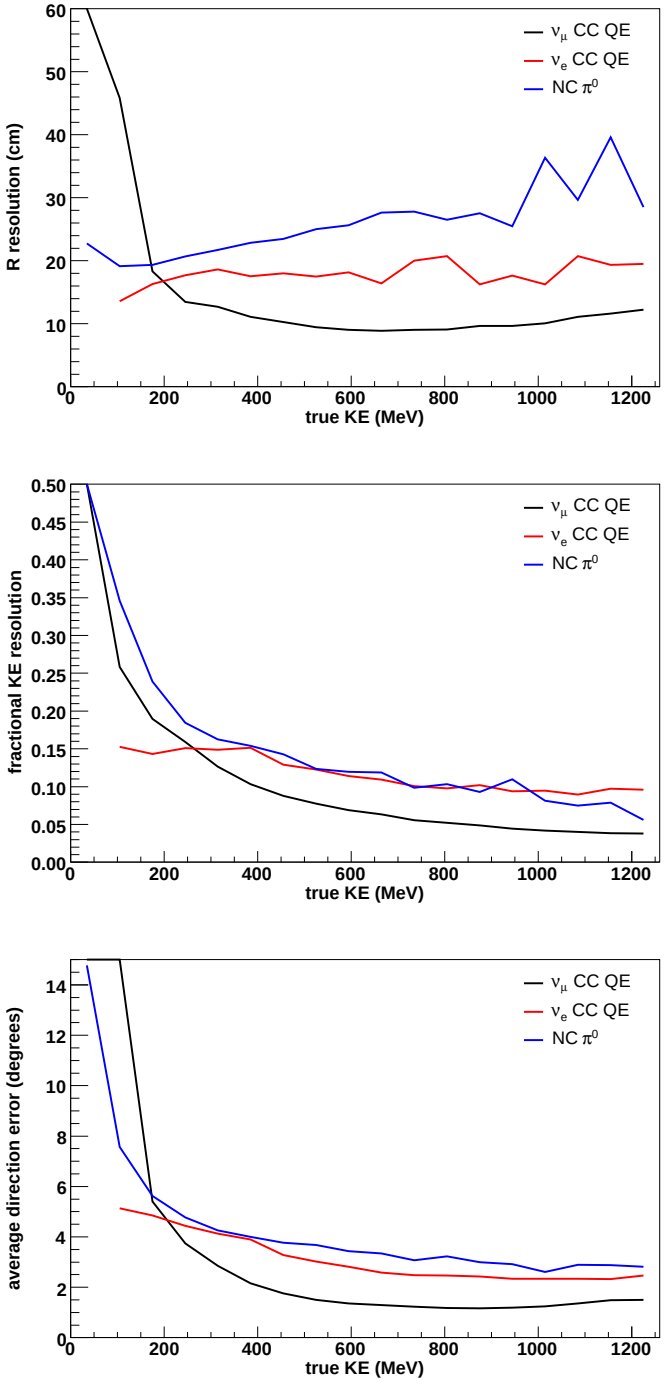

Fig. 18. Radius, kinetic energy, and direction resolutions. The jitter is due to limited test sample statistics in some regions. The ν_e curves cut off at 100 MeV due to an unrelated upstream cut on the Monte Carlo sample.

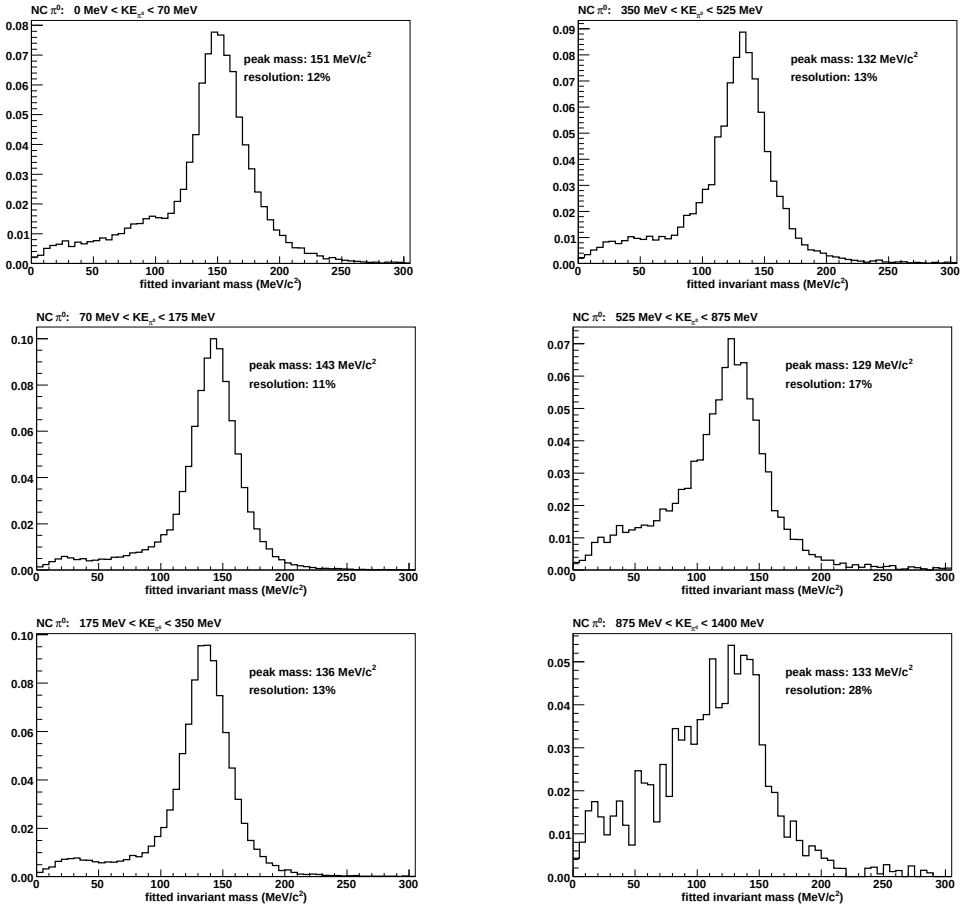

Fig. 19. Fitted invariant mass $M_{\gamma\gamma}$ for Monte Carlo-simulated NC π^0 events in three low energy regions (left) and three high energy regions (right).

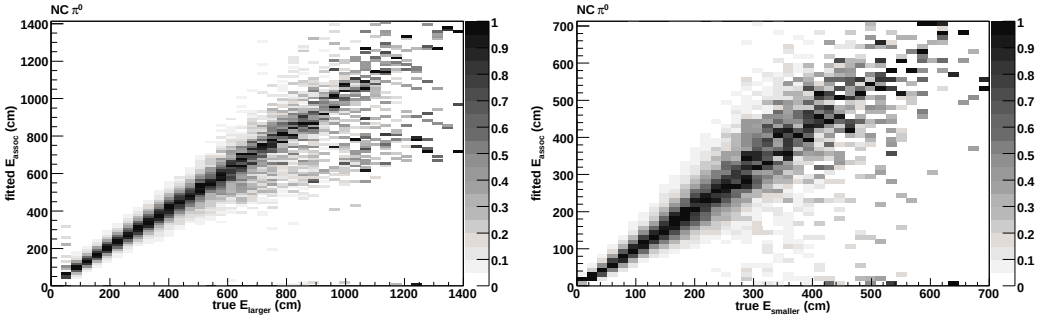

Fig. 20. Reconstructed γ energies in Monte Carlo-simulated NC π^0 events plotted against true values. The left (right) panel shows the higher (lower) energy γ from each event. The association of each fitted track to the underlying true γ's is, in general, ambiguous. For these plots, we choose the assignment that gives the smaller combined energy and direction discrepancy.

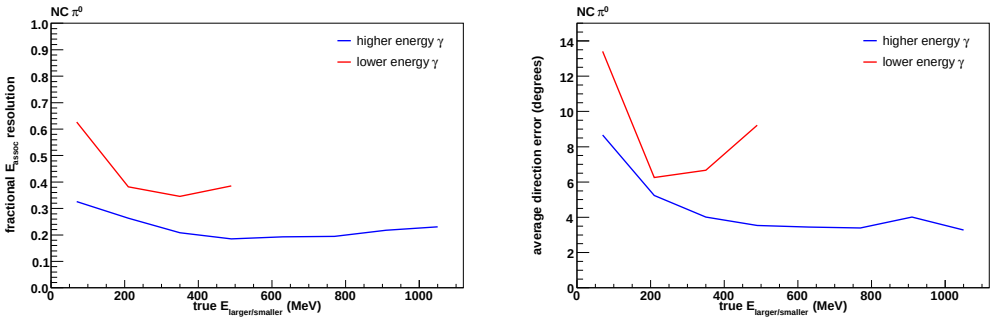

Fig. 21. Energy (left) and direction (right) resolutions for γ's in simulated NC π^0 events. For a given true γ energy, a dominant γ has expectedly better resolution than does a subordinate γ.

8 Particle Identification

The maximum likelihoods returned from each fit can be used for hypothesis testing. The two quantities

$$\mathcal{R}_{e/\mu} \equiv \log \frac{\mathcal{L}_e}{\mathcal{L}_\mu} = \log \mathcal{L}_e - \log \mathcal{L}_\mu \tag{37}$$

and

$$\mathcal{R}_{e/\pi} \equiv \log \frac{\mathcal{L}_e}{\mathcal{L}_\pi} = \log \mathcal{L}_e - \log \mathcal{L}_\pi \tag{38}$$

are used to determine for a given event whether the electron hypothesis is preferred over the muon and π^0 hypotheses. In these expressions, $\mathcal{L}_e$, $\mathcal{L}_\mu$ and $\mathcal{L}_\pi$ are the maximized likelihoods returned by the electron, muon, and (fixed-mass) two-track fits, respectively.

8.1 Electron/Muon separation

Figure 22 shows for simulated ν_e CCQE events and ν_μ CCQE events the distribution of $\mathcal{R}_{e/\mu}$ as a function of the energy reconstructed by the electron fit. For each sample, the events have been subject to a set of preselection criteria used in the ν_e appearance analysis. These require that there is only one time-cluster of PMT hits in the event, eliminating obvious ν_μ CC events that produce decay electrons. Cosmic backgrounds are eliminated by requiring the minimum number of PMT hits in the main region to be >200 and the number of veto hits to be <6. Furthermore, the average time of hits is required to lie within the expected beam delivery window from the Booster.

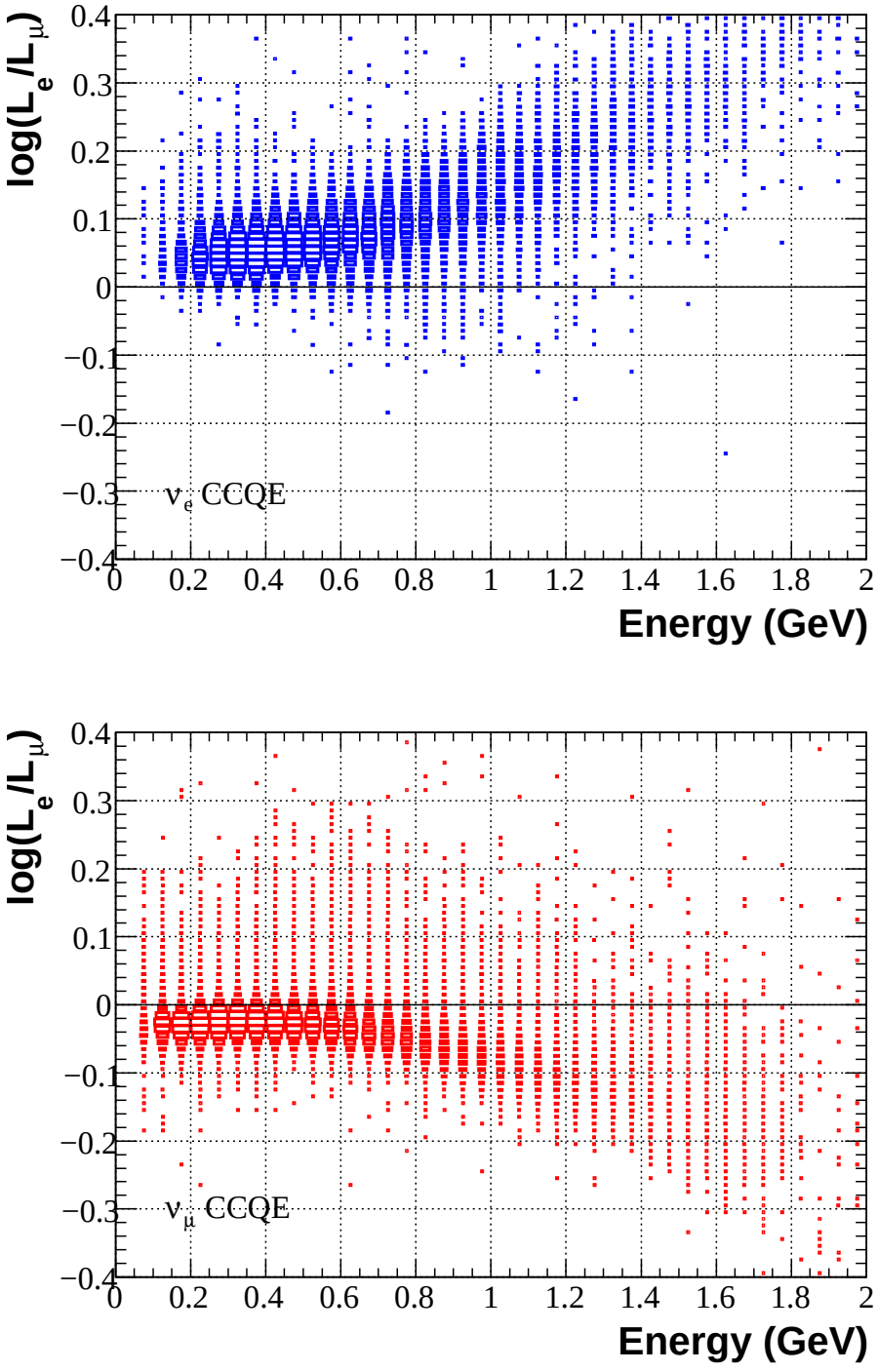

Fig. 22. Distribution of $\mathcal{R}_{e/\mu} = \log(\mathcal{L}_e/\mathcal{L}_\mu)$ for Monte Carlo simulated ν_e CCQE events (top) and ν_μ CCQE events (bottom) as a function of reconstructed energy (from the electron hypothesis fit).

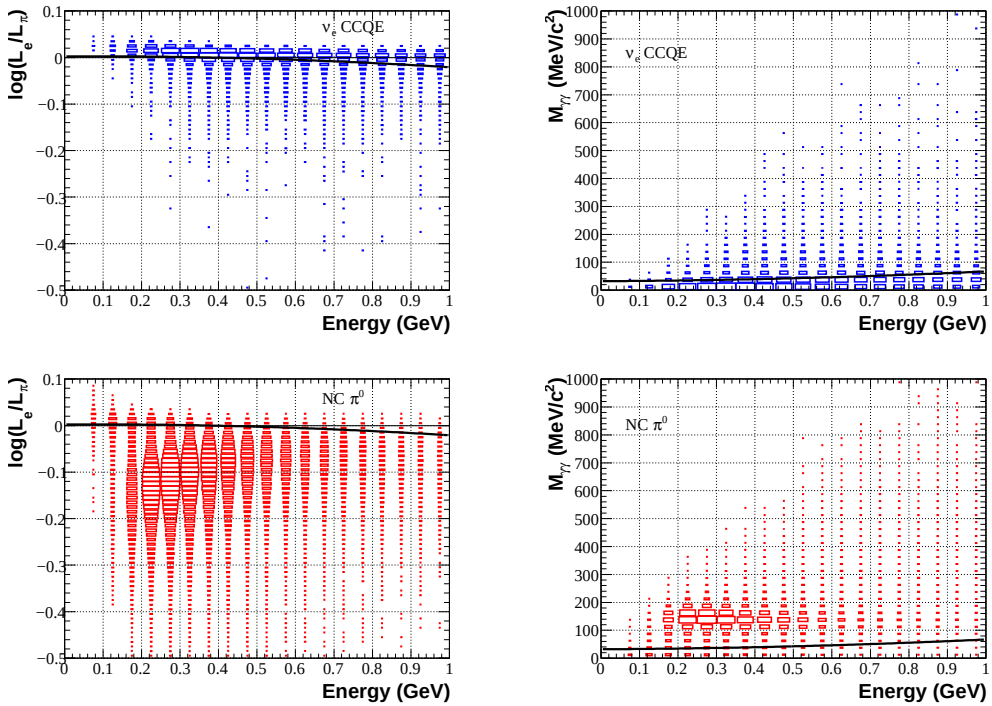

Fig. 23. Left: Distribution of $\mathcal{R}_{e/\pi} = \log(\mathcal{L}_e/\mathcal{L}_\pi)$ for Monte Carlo simulated ν_e CCQE events (top) and NC π^0 events (bottom). Right: same for $M_{\gamma\gamma}$. The black lines on each distribution show the selection used in the MiniBooNE $\nu_\mu \rightarrow \nu_e$ oscillation analysis [4].

For the ν_e CCQE events, $\mathcal{R}_{e/\mu}$ tends to take positive values, indicating that the fit to the event with the electron hypothesis is favored over the muon hypothesis. Likewise, $\mathcal{R}_{e/\mu}$ tends to be negative for the ν_μ CCQE events, indicating that the muon hypothesis fits these events better than the electron hypothesis. At high energies (>1 GeV), the electron/muon separation is aided by the fact that the muon pathlength grows approximately linearly with energy, while the electron shower grows more slowly.

8.2 Electron/π^0 separation

The left plots in Figure 23 show the distributions of $\mathcal{R}_{e/\pi}$ as a function of energy reconstructed by the single-track electron reconstruction for simulated ν_e CCQE events and NC π^0 events. The quantity $\mathcal{R}_{e/\pi}$ tests whether a given event fits better as a single electron track or as a π^0. Overall, electron/π^0 separation becomes more difficult at high energies as the energies of the decay photons in π^0 events become more asymmetric and the shower fluctuations in a single electron event get large enough to mimic the presence of a second photon. Also, π^0 events in which one of the two photons leaves the detector unconverted present an essentially irreducible background at low energy.

Another quantity that can be used for electron/π^0 separation is $M_{\gamma\gamma}$, the invariant mass of the two photons returned by the free-mass two-track fit. The NC π^0 events peak at the π^0 mass, whereas the ν_e CCQE events peak toward zero. This is seen in the right plots in Figure 23, where the $M_{\gamma\gamma}$ distribution is shown as a function of reconstructed energy. The use of the true π^0 mass in the fit's seeding procedure (Section 6.2) induces no appreciable high-mass peak in the ν_e CCQE events while helping considerably in the correct identification of NC π^0 events.

The ν_e appearance analysis at MiniBooNE utilizes both $M_{\gamma\gamma}$ and $\mathcal{R}_{e/\pi}$ to separate ν_e CCQE events from NC π^0.

8.3 Particle Identification Performance

The background rejection levels reachable with these fitter-derived quantities are shown for Monte Carlo events in Figures 24 and 25. Figure 24 shows the efficiency, after fiducial volume and cosmic rejection cuts, for selecting signal ν_e CCQE events alongside the efficiency (or misidentification rate) for ν_μ CCQE events. Three prototype selections are used, each chosen to give an approximately fixed signal efficiency at all energies. Note that the misidentification rate is shown with a logarithmic vertical scale. Figure 25 shows similar plots for π^0 rejection, one for the $\mathcal{R}_{e/\pi}$ cut and one for the $M_{\gamma\gamma}$ cut.

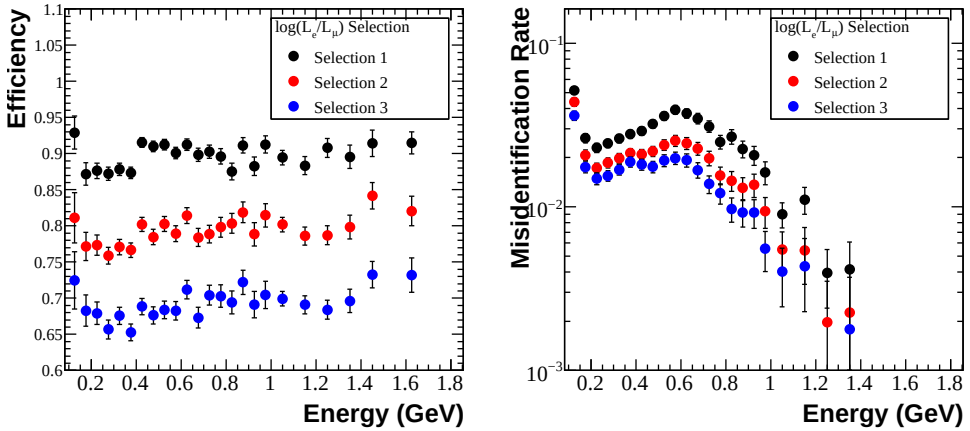

Fig. 24. Performance of three prototype $\mathcal{R}_{e/\mu}=\log(\mathcal{L}_e/\mathcal{L}_\mu)$ cuts for the ν_e selection. Left: Signal (ν_e CCQE) efficiency versus the electron reconstruction's energy. Right: Misidentification rate of ν_μ CCQE events versus the electron reconstruction's energy.

Misidentification rates of a few percent for ν_μ CCQE events are seen for signal efficiencies from $70-90\%$. (Note that ν_μ CCQE events are further reduced by about an order of magnitude by recognizing the delayed muon decay.) The two π^0 variables indicate that the misidentification rate increases with energy;

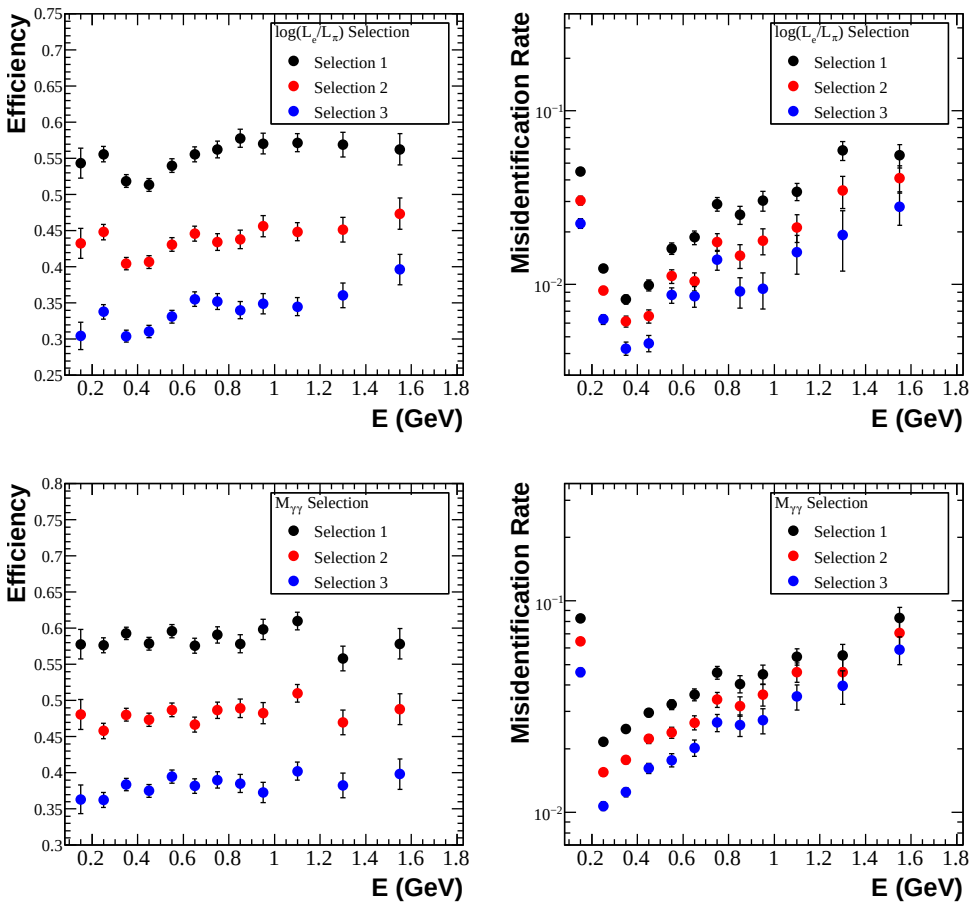

Fig. 25. Performance of three prototype $\mathcal{R}_{e/\pi}$=log($\mathcal{L}_e/\mathcal{L}_\pi$) (top) and $M_{\gamma\gamma}$ (bottom) cuts for the ν_e selection. Left: Signal (ν_e CCQE) efficiency versus the electron reconstruction's energy. Right: Misidentification rate of NC π^0 events versus the electron reconstruction's energy.

this is expected, since the two decay photons become highly asymmetric energetically and since the Cherenkov rings can achieve a greater degree of overlap. The rise in the π^0 misidentification rate at low energies is due to events with an escaping photon (corresponding to the cluster of π^0 events with positive $\mathcal{R}_{e/\pi}$ at 50−200 MeV in Figure 23). Between 200 MeV and 600 MeV, where most of the NC π^0 events lie, a misidentification rate of less than 1 percent can be achieved with 40% efficiency using only the $\mathcal{R}_{e/\pi}$ cut, while rates of a few percent can be achieved with only the $M_{\gamma\gamma}$ cuts. In the MiniBooNE $\nu_\mu \rightarrow \nu_e$ oscillation analysis, 50% signal efficiency and 1% π^0 misidentification rate are obtained by requiring both $M_{\gamma\gamma} < M^{\max}_{\gamma\gamma}(E)$ and $\mathcal{R}_{e/\pi} > \mathcal{R}^{\max}_{e/\pi}(E)$, where the "max" functions are quadratic in the energy E obtained from the single-track electron fit. These quadratic selection boundaries are indicated by the solid black lines in Figure 23. These cuts, along with an analogous $\mathcal{R}_{e/\mu}$ cut, were tuned to optimize sensitivity to LSND-like oscillations.

9 Summary

A maximum likelihood reconstruction algorithm is used at MiniBooNE to determine track parameters under electron, muon, and π^0 hypotheses. Underlying the likelihood is a track and detector model that calculates for a given parameter set the charge and time PDFs expected for each PMT, accounting for the spatially extended production of Cherenkov and scintillation light as well as the effects of indirect light from subsequent optical processes.

The maximized likelihoods for a given event obtained under different event hypotheses are used for event selection. In particular, the ratio of the likelihoods under the electron and muon models is used to suppress ν_μ CC events in the MiniBooNE $\nu_\mu \to \nu_e$ oscillation search; the ratio of the likelihoods under the electron and two-track π^0 fit (with fixed invariant mass) and the invariant mass obtained from the free-mass two-track fit are used to suppress NC π^0 events. While the reconstruction has been developed within the context of the MiniBooNE $\nu_\mu \to \nu_e$ oscillation search, other experiments employing similar Cherenkov detection techniques should be able to use the techniques discussed here.

References

[1a] K. S. Hirata *et al.* [Kamiokande-II Collaboration], Phys. Rev. Lett. **65**, 1301 (1990).

[1b] Y. Fukuda *et al.* [Kamiokande Collaboration], Phys. Lett. B **335**, 237 (1994).

[2a] S. Fukuda *et al.* [Super-Kamiokande Collaboration], Phys. Lett. B **539**, 179 (2002) [arXiv:hep-ex/0205075].

[2b] M. H. Ahn *et al.* [K2K Collaboration], Phys. Rev. Lett. **90**, 041801 (2003) [arXiv:hep-ex/0212007].

[2c] S. Yamamoto *et al.* [K2K Collaboration], Phys. Rev. Lett. **96**, 181801 (2006) [arXiv:hep-ex/0603004].

[3a] Q. R. Ahmad *et al.* [SNO Collaboration], Phys. Rev. Lett. **87**, 071301 (2001) [arXiv:nucl-ex/0106015].

[3b] Q. R. Ahmad *et al.* [SNO Collaboration], Phys. Rev. Lett. **89**, 011301 (2002) [arXiv:nucl-ex/0204008].

[3c] S. N. Ahmed *et al.* [SNO Collaboration], Phys. Rev. Lett. **92**, 181301 (2004) [arXiv:nucl-ex/0309004].

[4] A. A. Aguilar-Arevalo *et al.* [The MiniBooNE Collaboration], Phys. Rev. Lett. **98**, 231801 (2007) [arXiv:0704.1500 [hep-ex]].

[5] A. A. Aguilar-Arevalo *et al.* [The MiniBooNE Collaboration], Nucl. Instrum. Meth. A **599**, 28 (2009).

[6] S. J. Brice *et al.*, Nucl. Instrum. Meth. A **562**, 97 (2006).

[7] CERN Program Library Long Writeup W5013 (1993).

[8] D. Casper, Nucl. Phys. Proc. Suppl. **112**, 161 (2002) [arXiv:hep-ph/0208030].

[9] R. B. Patterson, Ph.D. thesis, Princeton University, FERMILAB-THESIS-2007-19 (2007)

[10] A. A. Aguilar-Arevalo *et al.* [MiniBooNE Collaboration], Phys. Lett. B **664**, 41 (2008) [arXiv:0803.3423 [hep-ex]].

[11] CERN Program Library Long Writeup D506